UNIVERSITY
OF MICHIGAN
JUL 13 1955
SOCIAL SCIENCE
LIBRARY

Journal of Experimental Psychology

BF
1
.J86

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Effect of Prior Patterns of Experience upon Strategies and Learning Sets: J. J. GOODNOW AND T. F. PETTIGREW 381

A Source of Decrement in Psychomotor Performance: J. A. ADAMS 390

Latent Extinction and the Reduction of Secondary Reward Value: H. MOLTZ 395

The Effects of Rest Periods on Interference of a Well-Established Habit: G. W. DOTEN 401

Classical and Instrumental Eyelid Conditioning: G. A. KIMBLE, L. I. MANN, AND R. H. DUFORT 407

An Evaluation of the Effect of Induced Tension on Performance: L. E. BOURNE, JR. 418

Stimulus Generalization Following Different Methods of Training: D. B. REINHOLD AND C. C. PERKINS, JR. 423

Stimulus Area, Stimulus Dispersion, Flash Duration, and the Scotopic Threshold: O. S. ADAMS, D. J. CHAMBLISS, AND A. J. RIOPELLE 428

The Influence of Overtone Structure on the Pitch of Complex Tones: W. H. LICHTE AND R. F. GRAY 431

Accuracy of Positioning Responses as a Function of Spring Loading in a Control: H. P. BAHRICK, W. F. BENNETT, AND P. M. FITTS 437

Reproduction of Simple Movements as a Function of Factors Influencing Proprioceptive Feedback: H. P. BAHRICK, P. M. FITTS, AND R. SCHNEIDER 445

American Psychological Association

Vol. 49 No. 6 June 1955

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1955 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Box 1557, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright, 1955, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 49, NO. 6 JUNE, 1955

EFFECT OF PRIOR PATTERNS OF EXPERIENCE UPON STRATEGIES AND LEARNING SETS[1]

JACQUELINE J. GOODNOW AND THOMAS F. PETTIGREW

Harvard University

This experiment was designed to investigate the effect of experience with one pattern of events upon the way *Ss* learn later patterns. Suppose one group of *Ss* begins by learning that there is a consistent pattern of reinforcement on a two-choice gambling device such that one side always pays off. We now shift them to a schedule in which payoff on the two sides is randomly arranged, averaging out to 50:50 over a series of trials. After exposure to this irregular schedule, a consistent pattern is reintroduced and the gambling device is again paying off regularly on one side, not necessarily the same side as before. Contrast this group of *Ss* with another treated identically in all respects save that they have never experienced an initial period of learning a consistent pattern of payoff. How quickly will the two groups learn the final consistent pattern? Will the effect depend upon whether or not *Ss* in the first group begin and end with the same side always paying off? Can differences in speed of learning the final pattern be related to differences in ways of responding to the period of randomness?

We know relatively little about response to changes in the frequency or probability of an event, as in a change from 100:0 payoff to 50:50 and then to 0:100 (**9, 15**). Recently there has been a trend towards looking at the problem in terms of "patterns of experience" (**12, 13**) and "patterns of reinforcement" (**7, 8, 9**) rather than exclusively in terms of simple frequency of reinforcement. Difficulty arises, however, when one comes to describe how it is that a previous pattern of experience affects later responses. Hilgard regards previous patterns of experience as affecting *Ss*' "expectations"; Hake, Grant, and Hornseth (**9**), following Harlow (**11**), suggest that "learning sets" are involved. Such terms as expectations and learning sets are notoriously difficult to define, and one would like to be able to describe more specifically the effects of previously experienced patterns. It seemed to us that an analysis of consistencies in the sequence of *S*'s responses might be rewarding (**2, 4, 10, 16, 19**). Suppose we find, for instance, that *S* consis-

[1] This research is one of a series of studies comprising the Cognition Project, Laboratory of Social Relations, Harvard University. The authors wish to express their appreciation of Dr. Jerome S. Bruner's advice.

tently repeats a choice of one alternative if it pays off but switches to the other alternative if he loses. Such a consistent way of deciding or responding we shall call a "strategy." We hope to be able to show that it is "strategies" that are affected by previously experienced patterns. We also hope to show that the speed of learning later patterns depends upon the appropriateness of the strategy previously developed. As one instance of this, we would like to show that experience with 100:0 is conducive to a strategy of staying with a winner and switching from a loser; that this strategy is maintained throughout the period of randomness; and that as a result the pattern of 100:0 is relearned very rapidly when it reappears.

Method

Apparatus.—The apparatus and general procedure have been described in detail elsewhere (6). Briefly, *S* plays a kind of two-armed bandit, with a choice on each trial between passing a left key or a right key. The *S* activates the machine by inserting a chip, and then presses the key which he thinks will pay off. If *S* selects the correct key, the apparatus drops a chip into a payoff slot. This is counted as a win. If no chip comes down, this is counted as a loss. The *S* is staked by *E* to 200 chips with which to play. At the end of the experiment, *S* is paid the difference between the number of times he has won and the number of times he has lost. The apparatus is operated by *E* from an adjoining room. The *S*s are led to believe, however, that the machine is automatic. They are given only one piece of information about the way the apparatus pays off. This is that one key and only one key will pay off on each bet.

Experimental design.—There were two general groups of *S*s, one given experience with consistency leading on to irregularity and then back to consistency again, and the other identical save that the initial experience of consistency does not occur.

Consistency was represented by two patterns of events—100% Left:0% Right and its reverse, 100% Right:0% Left. Irregularity was represented by two patterns of 50L:50R. As in many two-choice experiments, 50:50 means 50:50 occurrence of events rather than 50:50 reward. It means that the left key is scheduled to pay off on 50% of the trials while the right key is scheduled to pay off on the other 50% of the trials. One of the 50:50 patterns we shall call Short-Run 50:50, a series of 70 trials in which the 50:50 probabilities apply to every block of 10 trials. The other pattern we shall call Long-Run 50:50, a series of 70 trials in which the probabilities add up to 50:50 over the series of 70 trials but depart from 50:50 in each block of 10 trials. In the first block of 10 trials, the probabilities are 80L:20R; in the second block 40L:60R; and in subsequent blocks 60L:40R; 30L:70R, 70L:30R, 50L:50R, and 20L:80R. The terms "short-run" and "long-run" apply then in two senses. One is the number of trials needed before 50:50 probabilities apply. The other is the length of run of one event with which *S* is likely to meet, i.e., the number of times in a row that one key pays off.

The specific orders of events for different groups of *S*s are presented in Table 1. In general, the groups represent different combinations of initial consistency as 100L or 100R, irregularity as Long-Run 50:50 or Short-Run 50:50, and final consistency as 100L or 100R.

The *S*s were shifted from one pattern to another without any break. Wherever learning of a 100:0 pattern was involved, *S*s were given as many trials as they required to reach a criterion of 15 consecutive errorless trials. For 50:50 patterns, a standard series of 70 trials for each pattern was used, the same schedule being given to each *S* within a group.

Subjects.—The *S*s were Harvard undergraduates, predominantly freshmen. Within the limitation that we first filled all groups in which Long-Run 50:50 was involved, *S*s were assigned to groups in order of their appearance until there were ten *S*s in each group. The *S*s received a base pay of 75 cents in addition to any winnings on the bandit.

Results

The results can be most easily presented if we consider first the reasons for selecting 100:0 and 50:50 patterns as appropriate for showing the development and sustaining of strategies and their effect upon speed of later learning.

From an analysis of experiments where the pattern of events was changed after experience with a 100:0 pattern (**5, 9, 15**), it seemed that the effect of 100:0

reinforcement was to lead *S*s into what can be called a "win-stay; lose-shift" strategy or procedure. As long as an alternative was paying off, *S*s seemed to be repeating their choices of it, but shortly after it ceased to pay off, they would switch to the other alternative and stay with it for as long as it paid off. Now if this is the effect of experience with 100:0, then later behavior will not be differentially affected by whether the initial experience is one of 100L:0R or 100R:0L, even though the frequency of reinforcement of the two alternatives and presumably the response strengths of choosing left and right are different. Furthermore, the procedure of staying with the winning alternative should be sustained by Long-Run 50:50 where the long runs of either event and the frequent shifts from long runs of left to long runs of right make the strategy of staying with the winning alternative quite appropriate. After 100:0 and Long-Run 50:50, therefore, *S*s should be in an excellent position for relearning 100:0, and whether the initial and the final 100:0 are identicals or opposites should be of no importance.

In contrast, Short-Run 50:50 would not be expected to sustain the strategy of win-stay, lose-shift. One reason is that such a strategy leads to a considerable number of losses where the runs are short. Another way of looking at it is that Short-Run 50:50 is conducive to a sharply different strategy. From previous experiments, we know that *S*s given a Short-Run 50:50 schedule (this is the usual type of schedule in 50:50 studies), develop what can be called a "win-?; lose-stay" strategy (**14, 19**). The *S*s try to operate on the basis of the probable length of a run of events of payoffs. If they choose an alternative and it pays off, they may or may not switch to the other alternative depending on whether or not they consider that this is the end of the run. If they switch and lose, then they persist with the same choice because they consider that the run of payoffs on the other alternative is even more likely to be over. We would expect, therefore, that *S*s who start with 100:0 and then are shifted to Short-Run 50:50 would develop a mixture of both the strategies we have described and should be at no special advantage when it comes to relearning 100:0. Whether they should be at a special disadvantage, we did not know.

Within this framework, let us consider first *differences between groups*

TABLE 1

ORDER OF EVENTS FOR DIFFERENT GROUPS, AND MEDIAN TRIALS TO CRITERION OF 15 CONSECUTIVE ERRORLESS TRIALS IN INITIAL AND FINAL PHASES OF THE LEARNING SCHEDULE

Group	Order of Events*			Median Trials to Criterion†	
	Initial	Interpolated	Final	Initial	Final
1	A. 100L	Long-Run 50:50	100L	9	2
	B. 100R	Long-Run 50:50	100L	7	1.5
2	A. 100R	Long-Run 50:50	100R	8	2
	B. 100L	Long-Run 50:50	100R	8.5	2.5
3	——	Long-Run 50:50	100L	——	16
4	——	Long-Run 50:50	100R	——	13
5	A. 100L	Short-Run 50:50	100L	7	5.5
	B. 100R	Short-Run 50:50	100L	6	5.5
6	——	Short-Run 50:50	100L	——	10.5
7	——	Short-Run 50:50	100R	——	9.5

* As indicated, we did not run all eight of the groups that the design ideally calls for. This is because we first completed that half of the design which involved Long-Run 50:50. It was then clear that we could eliminate from the design the Short-Run 50:50 equivalents of Group 2.

† $N = 10$ each group.

in speed of learning a final 100:0 pattern. Median scores for all groups are presented in Table 1. As a summary statement before considering these scores in detail, the following points can be made. The speed with which 100:0 is relearned after initial 100:0 and Long-Run 50:50 is striking. This speed is not affected by whether or not initial and final 100:0 take the same form, e.g., both 100L:0R or one 100L:0R and the other 100R:0L. In contrast, initial 100:0 when followed by Short-Run 50:50 seems to leave *S*s little better off when it comes to relearning 100:0 than those *S*s with no initial experience of 100:0. There are also differences between the effect of Short-Run and Long-Run 50:50 when these are preceded by no initial training; the Long-Run 50:50 groups take longer here to learn final 100:0 than do the Short-Run 50:50 groups.

The questions listed below take up these points in more detail. The order in which the questions are dealt with is partly a matter of convenience rather than importance, since we wish to ask questions in such an order that we can increase the size of the sample by combining groups between which no significant difference is found.

Question 1.—Does it matter whether *S* starts and ends with the same form of consistency, whatever the interpolated form of 50:50 is? Does, for instance, initial training with 100L lead to any faster learning of final 100L than does initial training with 100R?

For the orders of experience used in the present experiment, it makes no difference whether the initial and final 100:0 are identicals or opposites. When it comes to learning 100L in the final phase, for instance, *S*s who have initially learned 100R learn just as quickly as *S*s who have initially learned 100L. This fits our general hypothesis that it is the patterning rather than the frequency of reinforcement that is important.

The statements above are based on comparisons within Groups 1, 2, and 5. Within each of these groups, half of the *S*s (subgroups A) start and end with the same form of 100:0, e.g., both are 100L; the other half of the *S*s (subgroups B) start and end with reverse forms of 100:0, e.g., they start with 100R and end with 100L. An inspection of Table 1 shows differences between the halves of the groups in the order of only 1 or .5 in the median trials taken to learn final 100:0. For more detailed comparison we used the Mosteller-Bush variation of the Mann-Whitney test (**17**), and obtained *z* scores of .34, .19, and .04 for Groups 1, 2, and 5, respectively. A *z* of 1.96 is regarded as grounds for rejecting the null hypothesis. Further reference to the *z* scores will be with reference to values obtained by the Mosteller-Bush form of the Mann-Whitney test. In the present context, the values of *z* obtained certainly substantiate the lack of difference between subgroups suggested by the medians in Table 1. Initial experience with 100L is as useful as experience with 100R for later learning of 100R. The question now arises of how useful initial experience with 100:0 is.

Question 2.—Is the final learning of 100:0 affected by previous experience with 100:0 as compared with no initial experience of 100:0?

This is one of the most important questions we are asking and the results are unequivocal. Final learning of 100:0 is definitely faster after initial experience of 100:0 or 0:100 than it is after no such experience, even though the initial and final experience of consistency are separated by 70 trials of 50:50 payoff. The difference is especially marked where Long-Run 50:50 is involved, *S*s with initial experience of 100:0 taking a median of only 2 trials to learn final 100:0 and *S*s without experience of 100:0 taking a median of 13.5 trials.

To determine the effects of initial experience of 100:0 as compared with no initial experience of 100:0, many comparisons could be made. However, we followed the simpler procedure of establishing first that there was no significant difference in the speed of

learning final 100L and learning final 100R,[2] and then pooled Groups 1 and 2, 3 and 4, and 6 and 7. We found only slight differences between learning 100L and learning 100R, generally in favor of faster learning of 100R, and none of the differences is statistically significant, the highest value of z being 1.06.

When we compare Groups 1 and 2 (initial experience of 100:0 and Long-Run 50:50) with Groups 3 and 4 (no such initial experience), we find that the median trials for final learning of 100:0 are 2 and 13.5 respectively. With the Mann-Whitney test, z is 4.82. When we compare Group 5 (initial experience of 100:0 and Short-Run 50:50) with Groups 6 and 7 (no such initial experience), the median trials for final learning of 100:0 are 5.5 and 9.5, respectively, and the value of z is 2.55. Clearly initial experience with 100:0 is an advantage in final learning of 100:0, and the advantage is greater with Long-Run 50:50 than with Short-Run 50:50.

Question 3.—Does experience with Long-Run 50:50 as against Short-Run 50:50 have a differential effect on the final learning of 100:0?

Long-Run 50:50 and Short-Run 50:50 do have differential effects upon the speed of learning final 100:0 but the direction of the difference depends on whether *S* experiences the 50:50 pattern after experience with 100:0 or at the very beginning of the learning schedule. Where the 50:50 patterns occur after experience with 100:0, Long-Run 50:50 leads to faster learning of final 100:0 than does Short-Run 50:50. Where, however, there has been no initial experience of 100:0, the findings are reversed, and Short-Run 50:50 leads to faster learning of final 100:0. This reversal is a reasonable result and is in line with arguments on resistance to extinction being a function of the discriminability of the difference between the new and the preceding pattern of events.

To test for differential effects where a 50:50 pattern occurs after initial experience with 100:0, we compared Group 1 (Long-Run 50:50) with Group 5 (Short-Run 50:50). The result is in favor of faster learning of 100:0 after Long-Run 50:50. Median trials to criterion are 2 and 5.5, respectively; the value of z is 4.67. To test for differential effects where there has been no initial experience of 100:0, we compared Groups 3 and 4 (Long-Run 50:50) with Groups 6 and 7 (Short-Run 50:50). The result is in favor of faster learning after Short-Run 50:50. Median trials to criterion are 13.5 and 9.5, respectively; the value of z is 2.08.

Question 4.—Are there any differences between learning 100:0 in the initial phase and learning 100:0 after different forms of previous experience?

Differences in speed of initial and final learning vary with the kind of experience which precedes the final 100:0. Where *S*s have learned 100:0 or 0:100 and have then experienced Long-Run 50:50, final learning of 100:0 is very much faster. Where *S*s have learned 100:0 or 0:100 and have then experienced Short-Run 50:50, final learning of 100:0 is no faster than initial learning of 100:0, even though relearning is involved and some savings might be expected. Where *S*s have not had initial learning of 100:0, they take longer to learn final 100:0 than *S*s normally take for initial learning of 100:0, the difference being especially marked after Long-Run 50:50.

The statements above are based on the following comparisons. For Groups 1 and 2, which both have Long-Run 50:50 interpolated between initial and final learning of 100:0, we compare *S*s' final

[2] We had trained some groups to learn final 100L and some groups to learn final 100R predominantly as a control against any interaction effects between the form of the final learning and any inadvertent patterning in the experimental schedule towards the end of the 50:50 period. For much the same reason we checked the effect of having especially long runs of payoff on one side at the beginning and end of the Long-Run 50:50 period. We checked this by running a control group where the Long-Run 50:50 phase started with a 60:40 rather than a 80:20 block of 10 trials and ended with the same block of 50:50 as in the Short-Run 50:50. The speed of learning final 100:0 was not affected by this rearrangement of the Long-Run 50:50 schedule.

learning against their own performance on initial learning. The difference is very much in favor of faster final learning of 100:0. Median trials to criterion are 2 and 7.5, respectively; the value of z is 6.52. The same kind of comparison is made for Group 5, which has interpolated Short-Run 50:50. Here we find no difference between the speed of initial and final learning of 100:0. Median trials are 6.5 and 5.5, respectively; z is .55. For groups with no initial experience of 100:0, we cannot compare groups against themselves. We use as a best estimate the initial learning of all other groups ($N = 60$). Comparing Groups 3 and 4 (Long-Run 50:50) and Groups 6 and 7 (Short-Run 50:50) against this standard, we find that final learning in both these groups takes longer than initial learning of 100:0 normally takes. Median trials to criterion are 13.5 and 9.5, respectively, as against a median of 7.5 in the standard groups; the values of z are 6.23 and 3.25, respectively.

Differences in strategies shown during the 50:50 period.—It will be recalled that we hoped to show not only that differences existed in speed of final learning, but also that these differences could be related to consistent sequential response patterns during the 50:50 period. We turn now to an analysis of the ways in which *S*s decided between alternatives, as indicated by differences in the way *S*s respond to wins and losses on the previous choice.[3]

The method for this analysis is as follows. We distinguish between four kinds of preceding trial in terms of the choice made (left or right key) and its outcome of win or loss. We then determine for each group the number of times that each kind of trial is followed by a choice of left and the number of times it is followed by a choice of right. We may find, for instance, that in Group 1 choices on trials immediately after a successful choice of left consist of

81% left key choices:
19% right key choices.

This ratio we compare with the over-all incidence of left and right choices in Group 1. We find that total choices are made up of

52% left key choices:
48% right key choices.

Clearly, *S*s in Group 1 are persisting with a choice of left, after a win on left, to a considerably greater extent than one would expect them to on the basis of their over-all choices of left. If we find the same tendency to persist with a choice of right when it pays off, we can argue that *S* has a general tendency to "stick to a winner." Our next step is to argue that it is this "strategy" of sticking with a winner and not persisting with a loser that enables Group 1 to respond so appropriately to a new pattern of final 100:0.

The incidence in all groups of left and right choices following win or loss on the previous choice is given in Table 2. Groups starting with 100:0 and then experiencing Long-Run 50:50 (Groups 1 and 2) show a marked tendency to repeat a previously successful choice. The ratio is 3 or 4 to 1 in favor of staying with a successful choice. This kind of strategy seems highly appropriate for the quick relearning of 100:0 shown by Groups 1 and 2.

Groups starting with Long-Run 50:50 also show a strong tendency to repeat a previously correct choice. The ratio is 1.5 or 2 to 1 in favor of staying with a successful choice. In contrast with Groups 1 and 2, however, these *S*s take a long time to learn final 100:0. The difference between the groups can be explained most easily by a specific

[3] We have also checked for differences between groups in (*a*) the total choices of left and right during the 50:50 period, which might be expected to differ after 100R as compared with after 100L, and (*b*) the lengths of runs of answers, which might be expected to differ after initial 100:0 as compared with no such initial experience. Total choices of left and right are differentially affected by 100R and 100L only for the first 10 trials of the 50:50 period. Apparently, the effect of initial reinforcement fades quickly. Lengths of runs of answers are amazingly similar in all groups.

TABLE 2

INCIDENCE OF LEFT AND RIGHT CHOICES, AND PROPORTION OF LEFT CHOICE FOLLOWING FOUR KINDS OF PRECEDING TRIAL
(Data for Trials 11–70 of 50:50 period)

Group	Event on Preceding Trial and Subsequent Choice														
	Chose L and Won			Chose L and Lost			Chose R and Won			Chose R and Lost			Total Choices		
	L	R	%L	L	R	%L	L	R	%L	L	R	%L	L	R	%L
1 & 2 (100:0 & Long-Run 50:50)	462	110**	81%	374	329	53%	127	448**	22%	289	261	53%	1252	1148	52%
3 & 4 (Long-Run 50:50 but no initial 100:0)	204	91*	69%	193	160	55%	119	168*	41%	129	136	49%	645	555	54%
5 (100:0 & Short-Run 50:50)	173	109*	61%	209	119*	64%	104	165*	39%	126	195*	39%	612	588	51%
6 & 7 (Short-Run 50:50 but no initial 100:0)	172	140	55%	174	95*	65%	150	180	45%	89	200*	31%	585	615	49%

* Significant at 1% level.
** Significant at .1% level. Significance levels established with binomial probability paper (18).

example. Where there is a run of 5 payoffs on the left key, *S*s who have had initial experience of 100:0 (Groups 1 and 2) tend to stay with the left key until the run is completely over. Only when it ceases to pay off do they switch to the right key. In contrast, *S*s who are experiencing Long-Run 50:50 without initial experience of 100:0 tend to stay with the left key for perhaps three or four of its payoffs but then to switch before the run is completely over. They have learned to anticipate longer runs of payoffs but they have not learned to ride the whole run out. When final 100:0 appears, they tend to play the 100% side for a number of times in succession but then to switch away from it.

Of the groups experiencing Short-Run 50:50 (Groups 5, 6, and 7), let us first consider those groups which have no initial experience of 100:0 (Groups 6 and 7). These groups show a marked tendency to persist with an unsuccessful choice. This same tendency has been reported in other experiments using Short-Run 50:50 schedules (14). What happens is that *S*s operate on the basis of the probable length of a run of payoffs. When they consider, for example, that a run of left payoffs is about to cease, they switch to right. If right does not pay off, *S*s choose it again because they figure that the chances are now even better that the run on the left is over. It is not clear whether *S*s generally expect runs to be shorter than are found in the experimental schedule or whether they operate on the basis of a modal run. On either basis, the strategy of persisting with an unsuccessful choice is likely to lead, as it does, to slower learning of final 100:0.

What of the group of *S*s who experience initial 100:0 and then Short-Run 50:50 (Group 5)? From the initial 100:0, they should learn to persist in choosing the key which is paying off. From the Short-Run 50:50, they should learn to persist with the key which is not paying off. In general, they should show either a mixture of both strategies or else show a replacement of the win-stay strategy by a lose-stay strategy. It is the former—a mixture of both strategies —which we find at all stages of the 50:50 period for this group. Whatever happens, *S*s tend to repeat their previous choice. Whether this is a true mixture of strategies, confusion, or a new procedure in its own right is difficult to determine. Ideally, we could carry the analysis further and either check the previous two or three trials or use an autocorrelation technique, but we are not sure that this would be rewarding.

Judging from *Ss*' statements and the relief with which they greet final 100:0, confusion and conflict is our best guess. In any case, it seems quite plausible that these *Ss* should learn final 100:0 neither faster nor more slowly than they learn initial 100:0.

Discussion

We started this experiment with the hope of showing that experience with one pattern of events can have a predictable effect upon the learning of later patterns of events, and that the effect can be mediated by the kind of strategy induced by the patterns previously experienced and the appropriateness of this strategy for later learning. We feel that these points have been quite well established. On a more general level, the results point to the usefulness of considering the problem of response to change in event occurrence in terms of such broad units as patterns of experience and sequential response patterning or what we have termed "strategies."

The effect of strategies needs to be regarded as twofold. In some situations, we can ask simply whether or not the strategy developed by *S* is directly appropriate for ease of later learning. This is the main factor in the ease with which final 100:0 is learned after a win-stay strategy has been developed during initial 100:0 and has been sustained throughout Long-Run 50:50. In other situations we need to ask whether the strategy which *S* brings to the new learning is indirectly appropriate in that it throws the changed pattern of events into relief and makes it highly discriminable. This seems to be the case in the faster learning of 100:0 after Short-Run 50:50 than after Long-Run 50:50 (when *S* has had no prior experience with 100:0). It is not only that *Ss* in Short-Run 50:50 have never experienced anything like such a long run before. It is also that the strategy of persisting with a wrong choice leads to such consistent error in the final 100:0 stage that they cannot overlook the change in the pattern of events. We are currently carrying out further research to provide more evidence of this indirect effect of strategies upon later learning.

Even granted this twofold effect, the question arises of whether one can cover all the possible effects of patterns of experience with such a concept as strategy. Our answer to this is no. There are undoubtedly other factors operating, even in the present experiment, that are more in the nature of general "learning sets." We shall mention two of these which are of particular interest.

The first is a question of the extent to which a particular order of experience teaches one to keep on watching how events occur rather than to regard the problem as solved and requiring only the repetition of previous procedures. This kind of learning set is probably most affected by experience with a number of transitions or reversals (**9, 11**), or by a lower than satisfactory level of success with previous procedures, or by the presence of different but recurring contingencies among events (**10**).

The second is a question of the extent to which a particular order of experience teaches one to ride with the pattern of events and to build hypotheses only after observation, rather than to use the events simply as a way of checking hypotheses independently or prematurely conceived. Differences in this kind of learning set seem to be largely responsible for the considerable differences among *Ss* in speed of initial learning of 100:0. The range is from 0 to 23 trials. Similar differences in problem-solving sets have been noted on an historical level by Stevens (**20**) and on an experimental level by Adishashiah (**1**). They seem to deserve greater prominence in the field of learning, especially where the situation calls for an awareness of the pattern of events and a matching of responses to this pattern.

Summary

The *Ss* chose on each trial between betting on the left or the right key of a kind of slot machine. Probability of payoff on the two keys is varied by *E*. The major interest is in the speed with which three groups of *Ss* learn the final set of probabilities. One group of *Ss* (A) starts with the left key always paying off (100L:0R), is then

switched to a randomly arranged 50L:50R, and finally brought back to 100L:0R. A second group of *Ss* has the same experience except that in the initial phase 100R:0L is substituted for 100L:0R. A third group has no initial experience with 100:0 but starts off with 50:50.

The main results are as follows:

1. In the final phase, having had initial experience with 100:0 is an advantage over having no such experience, even though 70 trials of 50:50 intervene between initial and final learning. It makes no difference, for the final learning of 100L:0R, whether the initial experience is of 100L:0R or of 100R:0L.

2. The speed of final learning varies with the arrangement of the 50:50 phase. For example, if events in this phase are arranged with a predominace of long runs, final learning in Group A is extremely fast (median of 2 trials as against 5.5 trials when the 50:50 phase has a predominance of short runs).

An analysis of response sequences shows that both these results are related to the way prior experience establishes and sustains consistent patterns of choice or "strategies," e.g., in Group A fast final learning of 100:0 follows the use throughout the 50:50 phase of a strategy of staying with a winning key and shifting from a loser. It is suggested that an analysis of such "strategies" is of value in obtaining more specific descriptions of the effect of prior experience upon later learning.

References

1. Adishashiah, W. Experimental studies on the influence of perceptual and thought processes on human performance. Unpublished doctor's dissertation, Cambridge Univ., 1951.
2. Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1954, in press.
3. Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
4. Frick, F. C. The effect of anxiety—a problem in measurement. *J. comp. physiol. Psychol.*, 1953, **46**, 120–123.
5. Gengerelli, J. A. Preliminary experiments on the causal factors in animal learning. *J. comp. Psychol.*, 1928, **8**, 435–457.
6. Goodnow, J. J. Determinants of choice distributions in two-choice probability situations. *Amer. J. Psychol.*, 1955, **68**, 106–116.
7. Grant, D. A., Riopelle, A. T., & Hake, H. W. Resistance to extinction and pattern of reinforcement: I. Alternation of reinforcement and the conditioned eyeblink response. *J. exp. Psychol.*, 1950, **40**, 53–60.
8. Hake, H. W., & Grant, D. A. Resistance to extinction and the pattern of reinforcement: II. Effect of successive alternation of blocks of reinforced and unreinforced trials upon the conditioned eyelid response to light. *J. exp. Psychol.*, 1951, **41**, 216–220.
9. Hake, H. W., Grant, D. A., & Hornseth, J. P. Resistance to extinction and the pattern of reinforcement: III. The effect of trial patterning in verbal "conditioning." *J. exp. Psychol.*, 1951, **41**, 221–225.
10. Hake, H. W., & Hyman, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45**, 64–74.
11. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
12. Hilgard, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.
13. Hilgard, E. R. The role of learning in perception. In R. R. Blake and G. V. Ramsey (Eds.), *Perception—an approach to personality.* New York: Ronald Press, 1951. Pp. 95–120.
14. Jarvik, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, **41**, 291–297.
15. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.
16. Miller, G. A., & Frick, F. C. Statistical behavioristics and sequences of responses. *Psychol. Rev.*, 1949, **56**, 311–324.
17. Mosteller, F., & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology.* Cambridge, Mass: Addison-Wesley, 1954. Pp. 289–334.
18. Mosteller, F., & Tukey, J. W. The uses and usefulness of binomial probability paper. *J. Amer. statist. Ass.*, 1949, **44**, 174–212.
19. Senders, V. L., & Sowards, A. Analysis of response sequences in the setting of a psychophysical experiment. *Amer. J. Psychol.*, 1952, **65**, 358–374.
20. Stevens, S. S. Psychology: the propaedeutic science. *Philos. of Science*, 1936, **3**, 90–103.

(Received August 26, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

A SOURCE OF DECREMENT IN PSYCHOMOTOR PERFORMANCE [1]

JACK A. ADAMS [2]

Skill Components Research Laboratory, Air Force Personnel and Training Research Center

It would seem reasonable to presume that a single measure of performance commonly taken for a trial on a complex psychomotor task is actually a composite score representing *S* accomplishment in several component responses. An *S* on the Rotary Pursuit Test not only makes the circular tracking response with his arm and hand but, at the same time, executes responses of visual discrimination, visual pursuit, and postural adjustment which are less obvious than the arm-hand response but nevertheless intimately linked with measured tracking proficiency.

Grice and Reynolds (**4**) have used a hypothesis based on response complexity considerations to account for reminiscence in the bilateral transfer of the rotary pursuit response (**4, 6, 7, 9**). The hypothesis accounts for this reminiscence in terms of the dissipation of reactive inhibition over rest for these less obvious component responses and the resulting increase in their reaction potentiality. Thus, in a bilateral transfer reminiscence investigation, part of the decrement observed in an initial training session results from reactive inhibition generated through repetition of these less obvious component responses. Then in a subsequent rest period, reactive inhibition dissipates and these responses increase in strength. When practice is resumed with the opposite hand in the final training session, a rest group will have higher time-on-target scores than a no-rest group because of the increase in strength for component responses common to the participation of either body member. Supporting evidence for this hypothesis would require a demonstration that repetition of a component response common to either hand contributes to time-on-target decrement and that this decrement dissipates over interpolated rest. The data of the present experiment bear on this issue.

Research findings to be reported are concerned with decrement in the over-all measured response arising from repetition of a component response. The component response selected for manipulation has been termed the visual response. The Rotary Pursuit Test, which was used in this experiment, can be considered a tracking task where *S* must visually perceive a target, continually discriminate the discrepancy between stylus tip and target, and attempt to correct this discrepancy by making appropriate arm-hand movements. Since this visual response of pursuit and discrimination is critical to successful performance of the task, inhibition of it through repetition should be reflected as decrement in the measured time-on-target score. Further, interpolated rest should lead to an increase in visual response strength

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the author. They are not to be construed as reflecting the views or indorsement of the Department of the Air Force.

[2] Now at the Interceptor Pilot Research Laboratory, Tyndall Air Force Base, Florida.

and a consequent increase in the level of time-on-target.

Method

Apparatus.—A 15-unit Rotary Pursuit Test was used and has been described elsewhere (8, p. 329–398). Time-on-target scores were recorded in units of .001 min. on Standard Electric timers. At the beginning of each trial, the disks turned and a buzzer sounded for 1.5 sec. before the clocks could be activated. Trial periods and intertrial rest periods were cycled automatically by Hunter interval timers.

Subjects.—The *Ss* were 356 basic airmen trainees drawn from the population available at Lackland AFB, Texas. Subgroups of 10 to 15 *Ss* were randomly assigned to the five main groups in the experiment.

Procedure.—For all groups, trial length was 1 min. and intertrial rest was 10 sec. except before and after the interpolated period where intertrial rest was increased to 30 sec. for instructional purposes. The groups and their experimental treatments were as follows:

Group A was the principal experimental group. Repetition of the visual response was interpolated between an initial and final practice session on the Rotary Pursuit Test. If exercise of the visual response is a source of decrement, then this group should have a lower performance level in the final session than the control conditions. Each *S* in this group had a partner who practiced on the unit to his left. Ten initial trials were given. Following Trial 10, Group A *Ss* were instructed to lay aside their styluses and note the small box at the left hand side of each unit. These were dummy boxes (*Ss* were not informed of this) with the button of a plunger-type microswitch protruding from the center and a pair of wires leading to the underside of the table. They were then told to (*a*) observe the performance of their respective partners during the next series of trials and depress the button on the box whenever they judged them to be on the target, and (*b*) press the button with the hand opposite to the one which held the stylus. This latter procedure was used to minimize expenditure of effort for the arm used in making the rotary movement. Twelve trials were then given in which the partners continued practicing on the Rotary Pursuit Test and Group A *Ss* observed them and pressed the button whenever they judged them to be on target. Following these button-pressing trials, Group A *Ss* and their partners practiced as a common group for 10 trials.

Group B was a control group. A rest period of 14 min. and 50 sec. was interpolated between an initial and final practice session of ten trials each and it was taken in a central waiting room where *Ss* sat and smoked, chatted, etc.

Group C was a control to evaluate the effects of button pressing per se. The possibility was admitted that work or interference effects of the button-pressing activity of Group A could lead to decrement in the final session quite independently of exercising the visual response. This group was a control for this possibility in that it had the same button-pressing activity as Group A but in response to an auditory signal rather than to the stimulus of a partner's performance. An initial and a final practice session of ten trials each were given. Interpolated between these two sessions were 12 trials of button pressing to a clearly audible click. Time interval between clicks was either .50, 1.0, or 1.50 sec. and the sequence of these interclick intervals was unsystematic. Button pressing was with the hand opposite to that which held the stylus.

Group D was the same as Group B except that the rest was taken in the testing room. These *Ss* stood in front of the apparatus during the rest period and were not permitted to lean on the table, smoke, or talk. Their purpose was to serve as control for the possibility that any decrement observed in the final session for Group A could be attributed to general fatigue arising from merely standing during the rest period and not necessarily to exercise of the visual response.

Group E received the same treatment as Group A except that a 10-min. rest in the central waiting room was interpolated between the end of the button-pressing session and the beginning of the final session on the Rotary Pursuit Test. It was expected that work inhibition would develop through exercise of the visual response in the button-pressing trials in the same fashion as for Group A, but that it would dissipate over the 10-min. interpolated rest period. The final session for this group was therefore expected to be essentially at the same level as Group B which merely rested.

The data of men who served as partners for *Ss* in Groups A and E were not analyzed since their sole purpose was to provide stimuli during the 12 button-pressing trials. The *N* for Groups A, B, C, D, and E, were 75, 74, 71, 74, and 72, respectively.

Results

Figure 1 shows the mean performance curves for the five groups. All groups received identical treatment in the initial ten trials and all are essentially at the same level. Using

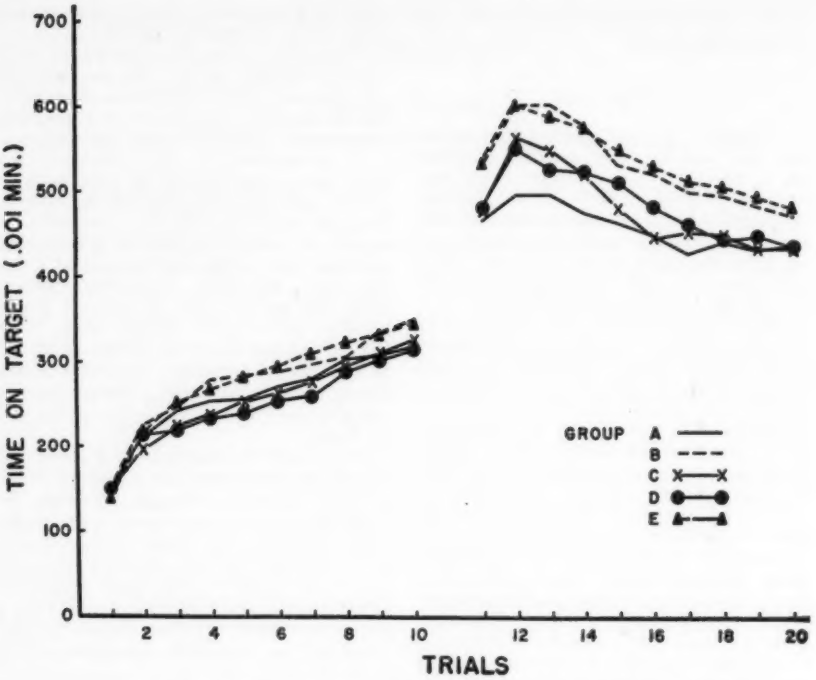

FIG. 1. Mean performance curves for the five groups. The break in the curves indicates the point of interpolated activity.

total score for the first ten trials as the measure, a simple analysis of variance test of the null hypothesis was made for the mean performance level in the initial session. An F of 1.11 for 4 and 361 df resulted in a retention of the hypothesis tested.

The final ten trials reflect the effect of the experimental manipulations and here the groups are differentiated. Group A, which had repetition of the visual response during the interpolated period, is well below the level of Group B which rested in the central waiting room. A test of the null hypothesis for the mean performance level in the final session was made for these two groups. Covariance analysis was used with the regression being total score for the final ten trials on total score for the first ten trials.[3] The ratio of the adjusted between-groups mean square to the adjusted within-groups mean square yielded an F of 16.88 for 1 and 146 df and permitted a rejection of the null hypothesis beyond the 1% level. The activity of Group A in the interpolated interval clearly served to depress performance below the rest condition.

Group E was to provide information on the stability of the decrement in Group A and it received the same experimental treatment as Group A except that a 10-min. rest was given between the end of button pressing and the beginning of the final practice

[3] These same measures were used in all covariance analyses made.

session. Inspection of Fig. 1 shows that Group A's decrement is transitory in nature since the rest period resulted in Group E's performance reminiscing to the level of Group B's. Covariance analysis was employed to test the null hypothesis for Groups A and E in the final session. The *F* of 23.06 for 1 and 144 *df* was significant beyond the 1% level and indicated that the additional 10-min. rest given Group E significantly affected performance. To test the hypothesis that Group B had reminisced to the same level as Group E, covariance analysis was used to evaluate the difference in their performance level in the final session. The *F* of .15 for 1 and 143 *df* sustained the hypothesis.

Groups C and D were control conditions included to more specifically identify sources of Group A's decrement. Group C was a check on whether button pressing per se led to work or interference factors that might effect performance in the final session, and Group D evaluated the possible decremental effects of merely standing during the interpolated period. The curves for these two groups in the final session have essentially the same level and trend and tend to fall above Group A but below Group B. The same level and trend suggests the hypothesis that button pressing to the auditory cue contributed no decrement beyond that generated by standing. Covariance analysis was used to test the null hypothesis for these two groups and an *F* of .68 for 1 and 142 *df* supported the hypothesis. To determine whether the standing response for Groups C and D placed them reliably below Group B which rested, the null hypothesis for these three groups was tested by covariance analysis. The obtained *F* of 2.57 for 2 and 215 *df* was significant between the 5 and 10% level and did not allow rejection of the hypothesis tested. It would appear that apparent decrement for Groups C and D in Fig. 1 is not supported statistically. To evaluate whether repetition of the visual response for Group A led to the development of decrement that could not be ascribed to the standing response, covariance analysis was used to test the null hypothesis for Groups A, C, and D. The *F* of 4.38 for 2 and 216 *df* was significant at the 2% level and permitted rejection of the hypothesis tested. It should be noted that the accumulation rate of inhibitory factors differs for Group A and Groups C and D since their substantial difference in the initial trials of the final session is absent at the end of the session.

A noteworthy finding is that repetition of the visual response for Group A serves to reduce noticeably the sharp initial rise in the final session performance curve which has come to be regarded as the warm-up segment (**1, 2, 5**).

Discussion

The results established that a source of decrement in the measured rotary pursuit response is repetition of the response *S* must make in visual pursuit of the target and in distinguishing the discrepancy between stylus tip and target. Further, evidence was presented to show that this decrement dissipates over rest. There was a suggestion that part of the decrement observed for the visual response could be ascribed to another component response of standing but this hypothesis could not be retained with confidence.

These findings make it plausible to assume that, in a study of bilateral transfer of the rotary pursuit response, repetition of the visual response in an initial practice session results in the generation of reactive inhibition. When practice in a final session is undertaken with the opposite hand, a group receiving intersession rest will show superiority over a no-rest group because inhibition

dissipation results in an increase in strength for the visual response which is a determiner of proficiency for either hand. The presumption would be that the visual response is only one of such component responses functioning in this fashion.

Repetition of the visual response in the interpolated period served to reduce the amount of warm-up decrement in the final session and this finding corroborates Irion's research (**5**) where interpolation of color naming reduced the amount of warm-up decrement in paired-associates learning. Ammons (**3**), however, failed to affect warm-up decrement by requiring ocular following of the target of the Rotary Pursuit Test during an interpolated session. The discrepancy between the present findings and Ammons' results is probably due to the procedures used in each case to insure *S* response. The method of evoking the visual response in the study reported here seems to guarantee *S* participation more effectively than Ammons' instructional approach. Irion's use of color naming in the verbal learning situation would seem to similarly guarantee *S* participation.

One shortcoming should be noted in the use of interpolated activity to lessen amount of warm-up decrement. Although button pressing reduced the amount of Group A's warm-up decrement, it also led to the accrual of reactive inhibition and a consequent lowering of general performance level in the final session. A worthwhile research effort would be to determine whether relatively small amounts of interpolated activity would materially reduce warm-up decrement but would generate negligible reactive inhibition. Future research will have to develop these relationships between warm-up decrement, decrement due to reactive inhibition, and the nature and amount of interpolated activity.

Summary

The single performance measure recorded for a trial on a psychomotor task can be regarded as a gross index of the level of several component responses constituting the response complex. Certain component responses are readily observable and unquestionably related to proficiency, but others such as visual discrimination and bodily adjustment are less obvious and their role in determining over-all performance level has been commonly ignored. Under investigation was the effect of component response repetition on decrement in the measure of over-all performance in the task.

The Rotary Pursuit Test was used. The component response of visual pursuit and discrimination was studied and was evoked by having *S* press a button each time he judged his partner to be on target. It was established that interpolation of visual response repetition led to significantly lower time-on-target scores in a final practice session. Control groups showed that this decrement could not be attributed to work or interference effects of button pressing per se or the general fatigue effects of standing. Another group demonstrated that this decrement dissipated over rest. Reduction in amount of warm-up decrement also resulted from repetition of the visual response. Implications for reminiscence in bilateral transfer were discussed.

References

1. Adams, J. A. Warm-up decrement in performance on the pursuit-rotor. *Amer. J. Psychol.*, 1952, **65**, 404–414.
2. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
3. Ammons, R. B. Effects of pre-practice activities on rotary pursuit performance. *J. exp. Psychol.*, 1951, **41**, 187–191.
4. Grice, G. R., & Reynolds, B. Effect of varying amounts of rest on conventional and bilateral transfer "reminiscence." *J. exp. Psychol.*, 1952, **44**, 247–252.
5. Irion, A. L. Retention and warming-up effects in paired-associates learning. *J. exp. Psychol.*, 1949, **39**, 669–675.
6. Irion, A. L., & Gustafson, L. M. "Reminiscence" in bilateral transfer. *J. exp. Psychol.*, 1952, **43**, 321–323.
7. Kimble, G. A. Transfer of work inhibition in motor learning. *J. exp. Psychol.*, 1952, **43**, 391–392.
8. Melton, A. W. (Ed.) *Apparatus tests.* AAF Aviation Psychology Program Research Report No. 4. Washington: U. S. Govt. Print. Off., 1947.
9. Rockway, M. R. Bilateral reminiscence in pursuit-rotor learning as a function of first-hand practice and length of rest. *J. exp. Psychol.*, 1953, **46**, 337–344.

(Received August 12, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

LATENT EXTINCTION AND THE REDUCTION OF SECONDARY REWARD VALUE [1]

HOWARD MOLTZ

University of Illinois[2]

Gleitman, Nachmias, and Neisser (**3**) posed the following question with reference to the S-R reinforcement theory of extinction: "Is performance a necessary condition for the extinction of a response" (**3**, p. 27)? Hull's theory of extinction would appear to require an affirmative answer, since reactive inhibition and conditioned inhibition could not develop unless the response, or some component of the response, was performed.

In order to test the hypothesis that a response can be weakened without that response being performed, Seward and Levy (**5**) trained rats to run down an elevated pathway to food. Following the last trial of learning, the experimental *S*s were placed on the goal platform for five 2-min. periods. During these periods the goal platform contained the food cup but did not contain food. Control *S*s were placed on a neutral platform for the same length of time. Seward and Levy concluded that the response tendency to run down the pathway was weakened without the response having been elicited, since the experimental *S*s reached an extinction criterion in significantly fewer trials than the control *S*s. The experimental *S*s also ran more slowly than the control *S*s during the extinction trials.

Bugelski, Coyer, and Rogers (**1**) took issue with the experimental design employed by Seward and Levy, pointing out that the experimental and control *S*s were detained on different platforms not only prior to extinction (pre-extinction experience), but also during the extinction trials. Bugelski, Coyer, and Rogers repeated the experiment with an additional control group that consisted of rats who were not given pre-extinction experience but who, like the experimental *S*s, were detained on the food platform between extinction trials. They failed to obtain evidence that pre-extinction experience on the food platform decreased significantly the number of trials to extinction.

Deese (**2**), however, found that pre-extinction experience was effective in reducing the strength of a response tendency. Deese used a U maze with one alley painted black that led to a black goal box; the other alley was painted white, and led to a white goal box. Food was placed in one of the goal boxes and all *S*s were trained to a position habit. Following the last learning trial, the experimental *S*s were given the opportunity to observe that food was no longer present in the goal box by being placed in the empty goal box for four 1-min. periods. Deese found that *S*s who were placed in the goal box made a smaller proportion of correct responses when subsequently run to extinction than control *S*s who were not "given the opportunity to inspect the empty goal box" (**2**, p. 366).

It is the present author's contention that although the results of Seward and Levy (**5**) and Deese (**2**) could not have been predicted from Hull's response-induced inhibition theory of extinction, these results can be interpreted in terms of Hull's general behavior theory (**4**). Accordingly, the purpose of the present experiment is to test the hypothesis that following training in a T maze, placing animals directly in the goal box without the presence of food (latent extinction) will reduce or eliminate the secondary reward value of cues present in the goal box, and that the effect of latent extinction in decreasing the number of subsequent trials to extinction will be dependent on this reduction or elimination of secondary reward value. If it is shown that one or more cues in the goal box had acquired secondary reward value, then the condition necessary for testing the experimental hypothesis would be satisfied.

Throughout the present study the

[1] An abstract of this paper was presented at the 1954 meeting of the American Psychological Association.

[2] Now at Brooklyn College.

term "latent extinction" will refer to the experimental procedure of placing animals in the correct goal box for four 1-min. periods without the presence of food. The term will *not* be used to refer to a process or to a state of the organism.

Method

Subjects and apparatus.—The *S*s were 60 experimentally naive, male albino rats of the Sprague-Dawley strain. The ages of *S*s ranged from 100 to 150 days at the beginning of the experiment.

The apparatus was a single-unit T maze and a straight alley. The starting box, arms, and stem of the maze were painted gray. The goal boxes were identical and were painted black. The roof of the maze was covered with .25-in. mesh hardware cloth. Black cloth curtains were placed in front of the doors to the goal boxes. Guillotine doors of light sheet metal were operated by *E*. There was a distance of 46 in. from the outer end of the left goal box to the outer end of the right goal box. The length of the stem was 21 in. The goal boxes were 10 in. long and 10 in. wide.

The straight alley was 24 in. long and 5 in. wide. The roof of the straight alley was covered with .25-in. mesh hardware cloth. The straight alley was also painted gray. The goal boxes were detachable and the same goal boxes were used on both the maze and the straight alley.

Several layers of cheesecloth surrounded completely the maze and the straight alley. The cheesecloth was attached to an unpainted wooden frame, 4 ft. wide and 3 ft. high. The frame was nailed to the table on which the maze and straight alley were placed. The use of the cheesecloth was intended to eliminate as many extra-maze cues as possible and to provide a nearly uniform perceptual field.

Preliminary training.—Prior to training, all *S*s were given seven days of habituation to the feeding regimen of the experiment. During these seven days, *S*s were fed an unlimited amount of wet mash once daily for 1 hr.

The preliminary training consisted of ten trials a day for three days in the straight alley. On seven of the ten trials of each day, *S* ran to food (consisting of pellets weighing approximately .150 gm.). The pellets were contained in a glass food cup of 2-in. diameter. The food cup was painted white. On the remaining three trials of each day the food cup was absent and *S* was not reinforced. The sequence of reinforced and nonreinforced trials within any day's run was determined randomly. The intertrial interval was always at least 10 min. with a modal interval of 12 min. Throughout the experiment the *S*s were under a 20–21-hr. hunger drive at the beginning of each day's run.

Learning series.—Beginning on Day 4, all *S*s were given ten trials a day in the maze for four days. The side of the maze in which any *S* was to receive food was determined randomly. For the remainder of the learning series food appeared only on that side. Runs to the correct side were always reinforced by the presence of food and by the presence of the food cup. The food cup was not present in the incorrect side of the maze. Responses to the correct or incorrect side of the maze did not involve any delay. Correction or retracing after an incorrect response was not allowed. All *S*s remained in the goal box for 45 sec. following either a correct or an incorrect response.

After every five trials of the learning series, the right goal box was changed to the left side of the maze and the left goal box was changed to the right side of the maze. Though the goal boxes were constructed to be identical, there might have been differences overlooked by *E* (e.g., odor, texture of the floor, etc.). Interchanging the goal boxes prevented any one goal box from being associated consistently with reward.

Latent extinction period.—Approximately 24 hr. after the last trial of the learning series, *S*s were assigned at random to three groups. Each group consisted of 20 *S*s. The *S*s of Group 1 and Group 2 (experimental *S*s) were placed directly in the goal box for four 1-min. periods without the presence of food. Each period was separated by approximately 10 min. For the *S*s of Group 1, the empty food cup was present in the correct goal box; for the *S*s of Group 2, the food cup was absent. The *S*s of Group 3 (control *S*s) were not placed in the goal box during the latent extinction period.

Extinction trials.—Immediately following latent extinction, response extinction was begun. Food was not present in the maze during the extinction trials. Each group was divided randomly into two subgroups. For the *S*s of one subgroup, the empty food cup was present in the previously correct goal box during the extinction trials; for the *S*s of the other subgroup, the food cup was absent during the extinction trials. Each subgroup consisted of 10 *S*s.

All *S*s were run until they reached the extinction criterion of two incorrect responses within any block of four successive trials. The extinction score of each *S* includes only the

number of correct responses made before reaching the extinction criterion.

During extinction, the intertrial interval was approximately 10 min. As in the learning series, correction or retracing after an incorrect response was not permitted. All *Ss* remained in the goal box for 10 sec. following a response either to the correct or the incorrect side of the maze.

Results and Discussion

Learning series.—No significant differences were found between the various groups in the number of correct responses made during the learning series.

Extinction series.—An analysis of variance and *t* tests using original extinction scores did not seem desirable since the requirement of homogeneity of variance could not be met. However, the transformation $\sqrt{X+.5}$ satisfied the requirement (6) and therefore all analyses were made using transformed data. Table 1 shows the means and *SD*'s of the raw data and of the transformed data for each group.

Table 2 shows the comparisons between the means of the transformed extinction scores of each group. Table 3 reveals the analysis of variance of the transformed extinction scores for the *Ss* of Groups 1 and 2.

In order to test the experimental hypothesis, it was first necessary to demonstrate that the conditions employed during preliminary training and during the learning series were adequate to establish the food cup as a secondary reinforcer. Table 1 shows that those *Ss* of Group 3 for whom the empty food cup was present during the extinction trials required a greater mean number of trials to reach the extinction criterion than those *Ss* of Group 3 for whom the food cup was absent during the extinction trials. The *t* shown in Table 2 indicates that the difference between the means of the transformed extinction scores of these subgroups is significant below the .01 level of confidence. There can be little doubt that the condition necessary for testing the experimental hypothesis was satisfied.

TABLE 1

Means and *SD*'s of Original and Transformed Data

Group	Extinction	Original Data		Transformed Data	
		Mean	*SD*	Mean	*SD*
1	Cup present	3.10	2.85	1.78	.70
	Cup absent	1.20	2.10	1.17	.61
2	Cup present	9.40	7.04	2.99	1.01
	Cup absent	1.50	1.51	1.32	.53
3	Cup present	9.30	6.46	2.98	.87
	Cup absent	3.10	3.11	1.73	.81

TABLE 2

Comparisons Between Means of Transformed Extinction Scores

Comparisons	*t*	*p*
Group 1, Group 2	2.94	<.01
Group 1, Group 3	3.16	<.01
Group 2, Group 3	.62	>.50
Group 1, cup present; Group 1, cup absent	1.84	<.08
Group 2, cup present; Group 2, cup absent	5.08	<.01
Group 3, cup present; Group 3, cup absent	3.16	<.01

TABLE 3

Summary of Analysis of Variance for *Ss* of Groups 1 and 2

Source	*df*	Mean Square	*F*
Group 1, Group 2	1	4.69	8.68**
Food cup present, absent, during extinction	1	13.03	24.13**
Group 1, Group 2 × presence and absence of food cup during extinction	1	2.85	5.28*
Within cells	36	.54	

* Significant at the .05 level.
** Significant at the .01 level.

In order to conclude that the secondary reward value of the food cup was either reduced or eliminated when the food cup was present during latent extinction, it is necessary to examine the interaction between the presence and absence of the food cup during latent extinction (Groups 1 and 2, respectively) and the presence and absence of the food cup during the extinction trials. Table 3 reveals that this interaction is significant below the .05 level of confidence. In addition, Table 1 shows that those *S*s of Group 2 for whom the food cup was present during the extinction trials required a greater mean number of trials to reach the extinction criterion than those *S*s of Group 2 for whom the food cup was absent during the extinction trials. The t reported in Table 2 indicates that the difference between the means of the transformed extinction scores of these subgroups is significant below the .01 level of confidence.

Table 1 also shows that for those *S*s of Group 1 for whom the food cup was present during the extinction trials, the mean number of trials to extinction was 3.10; for those *S*s of Group 1 for whom the food cup was absent during the extinction trials, the mean number of trials to extinction was 1.20. The t reported in Table 2 reveals that the difference between the means of the transformed extinction scores of these subgroups is not significant at the .05 level.

Since the interaction reported in Table 3 is significant and, in addition, since the difference between the subgroups of Group 1 is not significant while the difference between the subgroups of Group 2 is significant, it is concluded that the secondary reward value of the food cup was reduced significantly more for those *S*s for whom the food cup was present during latent extinction (Group 1) as compared with those *S*s for whom the food cup was absent during latent extinction (Group 2).

While the analyses indicated that the food cup had acquired secondary reward value and that this secondary reward value was reduced significantly more for those *S*s for whom the food cup was present during latent extinction as compared with those *S*s for whom the food cup was absent during latent extinction, several additional comparisons need be made before the experimental hypothesis is supported. If the effect of latent extinction in decreasing the number of subsequent trials to extinction is dependent on the reduction or elimination of secondary reward value, then those *S*s for whom the secondary reward value of the food cup was not significantly reduced (Group 2) should not differ from those *S*s who were not given latent extinction (Group 3). On the other hand, those *S*s for whom the secondary reward value of the food cup was significantly reduced (Group 1) should have required fewer trials to reach the extinction criterion than either the *S*s of Group 2 or the *S*s of Group 3. Analysis of the data supported these predictions. The mean number of trials to extinction for the *S*s of Group 2 was 5.45; the mean number of trials to extinction for the *S*s of Group 3 was 6.20. Table 2 shows that the difference between the means of the transformed extinction scores of these groups is *not* significant. Table 2 also shows that the *S*s of Group 1 required significantly fewer trials to reach the extinction criterion than either the *S*s of Group 2 ($p < .01$) or the *S*s of Group 3 ($p < .01$).

Additional evidence in support of the experimental hypothesis can be obtained by comparing those subgroups of Groups 1, 2, and 3 for which the food cup was absent during the extinction test trials. If the effect of latent extinction is dependent on the reduction or elimination of the secondary reward value of the food cup, then the latent extinction effect should tend to "disappear" when the

food cup is not present during the extinction test trials. That is, there should be no significant differences between those subgroups of Groups 1, 2, and 3 for which the food cup was absent during the extinction trials. Analysis of the data supported this prediction; the values of t obtained by comparing the means of the transformed extinction scores of these subgroups were all above the .05 level of confidence.

On the other hand, if the latent extinction effect is dependent on the reduction or elimination of secondary reward value of the food cup, then the effect of the food cup in maintaining the previously learned response against extinction should be dependent on whether or not it was present during the latent extinction period. If the food cup was present during latent extinction (Group 1), then its presence during the subsequent extinction test trials should have little or no effect in maintaining the previously learned response against extinction. However, if the food cup was absent during latent extinction (Group 2) or if latent extinction was not given (Group 3), the presence of the food cup during the extinction test trials should serve to increase the number of trials to extinction. Thus, those subgroups of Groups 2 and 3 for which the food cup was present during the extinction trials should not differ significantly from each other, while they both should differ significantly from the corresponding subgroup of Group 1. Analysis of the data revealed that this prediction was supported; the mean number of trials to extinction for that subgroup of Group 1 was 3.10, as compared with 9.40 ($t = 3.69$; $p < .01$) and 9.30 ($t = 3.50$; $p < .01$) for the corresponding subgroups of Groups 2 and 3, respectively. The latter two subgroups did not, of course, differ significantly from each other.

It is concluded that the experimental hypothesis,—that the effect of latent extinction in decreasing the number of subsequent trials to extinction would be dependent on the reduction or elimination of secondary reward value,—was supported. It therefore appears likely that the results obtained by Seward and Levy (5) and Deese (2) can be interpreted in terms of the reduction or elimination of the secondary reward value of those cues present in the goal box when the animal was placed in the goal box without the presence of food. Deese, for example, used a U maze with one alley painted black which led to a black goal box; the other alley was painted white and led to a white goal box. Since food appeared only in the black (or white) goal box during the learning series, the color of the correct goal box may be expected to have acquired secondary reward properties which were subsequently reduced or eliminated when the animal was placed in the goal box without the presence of food.

In the procedure used by Seward and Levy, the food cup that was present on the goal platform during the learning series was also present on the goal platform during latent extinction. If the food cup had acquired secondary reward value, then on the basis of the results of the present experiment, it appears reasonable to expect that the secondary reward value of the food cup was reduced or eliminated for those *S*s for whom the food cup was present on the goal platform during latent extinction. Since for these *S*s the food cup did not subsequently reinforce the response of running down the pathway to the goal platform, they required fewer trials to extinction than control *S*s who were not placed on the goal platform and for whom the secondary reward value of the food cup was not reduced or eliminated.

Summary

The purpose of the present experiment was to test the hypothesis that the effect of latent extinction in decreasing the number of trials to extinction would be dependent on the reduction or elimination of the secondary reward value of cues present in the goal box during the latent extinction period. Sixty male albino rats were used. Following preliminary training, all *S*s were given 40 trials in a single-unit T maze with food reward in one of the goal boxes. After the last trial of the learning series, the *S*s of Groups 1 and 2 were placed in the goal box for four

1-min. periods without the presence of food (latent extinction). In Group 1 the empty food cup was present in the goal box during latent extinction; in Group 2 the food cup was absent during latent extinction. Group 3 was not given latent extinction. Immediately following the last latent extinction period, response extinction was begun. Each group was divided into two subgroups. In one subgroup the empty food cup was present in the previously correct goal box during the extinction trials, in the other the food cup was absent.

The extinction scores of the subgroups of Group 3 indicated that the food cup had acquired secondary reward value. The extinction scores of the subgroups of Groups 1 and 2 indicated that the secondary reward value of the food cup was reduced significantly more for those *S*s for whom the food cup was present during latent extinction (Group 1) as compared with those *S*s for whom the food cup was absent during latent extinction (Group 2). No significant difference was found between the means of the transformed extinction scores of Group 2 as compared with Group 3. Since, in addition, Groups 2 and 3 required significantly *more* trials to extinction than Group 1, the experimental hypothesis was supported.

References

1. Bugelski, B. R., Coyer, R. A., & Rogers, W. A. A criticism of pre-acquisition and pre-extinction of expectancies. *J. exp. Psychol.*, 1952, **44**, 27–30.
2. Deese, J. The extinction of a discrimination without performance of a choice response. *J. comp. physiol. Psychol.*, 1951, **44**, 362–366.
3. Gleitman, H., Nachmias, J., & Neisser, U. The S-R reinforcement theory of extinction. *Psychol. Rev.*, 1954, **61**, 23–33.
4. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.
5. Seward, J. P., & Levy, N. Sign learning as a factor in extinction. *J. exp. Psychol.*, 1949, **39**, 660–668.
6. Snedecor, G. W. *Statistical methods.* Ames, Ia.: Iowa State Coll. Press, 1946.

(Received June 7, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

THE EFFECTS OF REST PERIODS ON INTERFERENCE OF A WELL-ESTABLISHED HABIT[1]

GEORGE W. DOTEN[2]

Northwestern University

Most studies dealing with transfer effects (including retroactive inhibition studies) in verbal learning have been concerned with habits acquired in the laboratory. That is, *Ss* are required to learn one task to a certain criterion of mastery and then are given a second task to learn. The net amount of transfer, either positive or negative, is measured by comparing the scores made in learning the second task to the same measures of a control group who did not learn the first task.

In the present study an attempt was made to produce negative transfer in a verbal habit of long standing. Specifically, this study was oriented towards the production of a decrement in the reading habit, and towards examining the influence of two variables related to this decrement.

In 1935, Stroop (3) published a report using materials which seem to be appropriate for studying decrements in the reading habit. Stroop constructed lists of color words printed in another color. For example, the word "red" was printed in blue, the word "green" in brown, and the word "brown" in red. He then had *Ss* learn to ignore what the word said and name the color of the word (naming habit). After eight days of practice, four trials each day on a 50-word list, *Ss* were instructed to read the words, now ignoring the color of the word. When the pre- and postreading performances were compared, Stroop found that interference was produced. That is, it took *Ss* longer to read the words following practice on naming the color of the word.

Basically, Stroop's materials were used in the present study. However, two additional variables were manipulated, namely, intertrial distribution on the naming task and the time interval between the naming task and the postreading task. The study was conducted as two separate experiments. Experiment I was designed to discover how distributed practice on the naming task effected the postreading. In this experiment, the reading task followed the last trial on the naming task almost immediately. In Exp. II, the same degrees of distribution were used as in Exp. I; however, the time interval between naming and postreading tasks was extended.

METHOD

Subjects.—A total of 90 *Ss*, 45 in each experiment, were used. The *Ss* were men and women volunteers from the introductory psychology course at Northwestern University.

Materials.—The materials consisted of a stop watch, color blindness test, and a white cardboard mounting the color words. The cardboard measured 12 × 14 in. and contained 16 lines of color words, nine per line, each line spaced ½ in. apart. Except for two restrictions, the color and order of the words was random. The two restrictions were: (*a*) no color word was printed in the color it represented, and (*b*) all colors appeared equally often, except yellow, which appeared one less time. Only five different color words and colors were used, namely, red, blue, green, yellow, and brown. Only one order of the words and colors was employed although three cardboard mountings

[1] This article is based on a thesis presented in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the Department of Psychology, Northwestern University. The writer is greatly indebted to Dr. B. J. Underwood for his encouragement and generous aid.

[2] Now at St. Lawrence University.

were constructed to avoid fading of colors as the experiment proceeded.

Conditions.—All *Ss* were given five trials on the reading task (reading the words; ignoring the color of the words). Following this, four days of practice, 15 trials per day, were given on the naming task (calling out the color of the words). Following this a posttest was given on the reading task to determine the effects of the four days of practice in naming.

The two experiments differed in only two ways. In Exp. I, 30 sec. elapsed between the last trial on the fourth day of color naming and the reading test; in Exp. II this interval was 24 hr. In Exp. I, 5 posttest reading trials were given; in Exp. II, 10 trials.

Within each experiment three groups of *Ss* received 0-, 30-, or 60-sec. rest between naming trials. All *Ss* were given 30-sec. rest between all reading trials.

A trial consisted of reading or naming 144 words or colors of the words. The response measure was the time per trial. The *Ss* were assigned to each group on the basis of their time score made on the first naming trial. In this way the groups were equated on initial skill on this task.

Procedure.—Each *S* that came into the laboratory was first given a color blindness test. The *S* was then given the cardboard with the 144 stimuli and instructed to read the words as fast as possible. At the end of each trial, he was told to turn the list over and wait for a ready signal to start the next trial. At the end of the fifth reading trial, *S* was told that he now must ignore the color word itself and name the color in which each color word was printed. An *S* in the 30-sec. or 60-sec. group was again told to turn the list over at the end of each trial and wait for a ready signal for the next trial. However an *S* in the 0-sec. group was instructed to go back to the beginning of the list immediately when he came to the end of the list. On Day 4 in Exp. I and Day 5 in Exp. II, *S* was again instructed to read the color words. All *Ss*, irrespective of group, were again instructed to turn the list over at the end of each trial and wait for a ready signal for the beginning of the next trial.

Results

Naming task.—Figure 1 illustrates the performance of the groups on the naming task for both experiments combined ($N = 30$ each point). However, Trials 1 and 15 were ana-

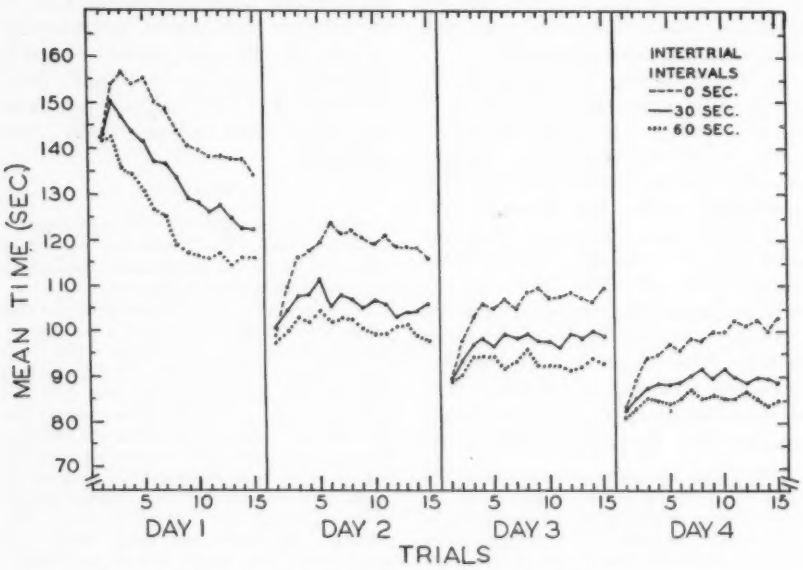

Fig. 1. Performance curves of each day on the naming task for each group for both experiments combined.

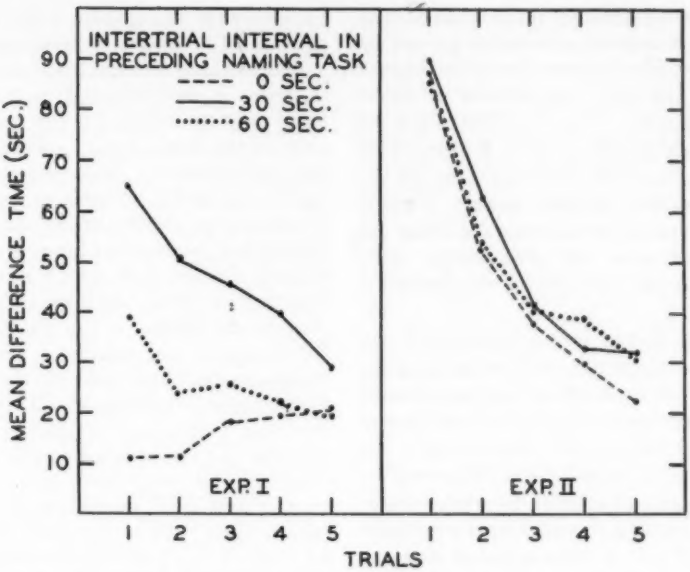

FIG. 2. Comparison between Exp. I and Exp. II of amount of interference on the reading task after 60 trials on the naming task.

lyzed statistically for each experiment separately. When the last trials of each day of Exp. I were analyzed by a technique suggested by Edwards (**1**, pp. 288–297), trial differences and an interaction, Trials × Methods, were found to be significant. Significant trial differences represented significant changes in learning from day to day, whereas the significant interaction effect indicated the approach in performance level of the 30-sec. group to that of the 60-sec. group. The overall difference among groups as to the effectiveness of the distribution variable was tested by means of analysis of covariance using the first naming trial on Day 1 (the matching score) as the covariate. This analysis was performed on the summed last trials of each group. An F of 5.08 ($.02 > p > .01$) was obtained. It can be seen from this analysis that the groups differ on this variable.

Analysis of the first trial of each day revealed only insignificant trial differences.

The same analyses were made on the first and last naming trials of each day in Exp. II. The results were the same as for Exp. I with one exception: the interaction effect of Trials × Methods was not significant.

From the results on the naming task (Fig. 1), it becomes apparent that while the groups differ in performance level, they do not differ in learning. That is, difference in performance within days is apparent as a function of intertrial rest, but all groups start at about the same level each day. In short, there is no evidence that the intertrial intervals used here produce permanent differences in naming habit.

Reading task.—Figure 2 and Table 1 present the performance of the groups on the reading task in both experi-

ments. Although there was no significant difference between groups on the reading pretest in either experiment the data here consist of difference scores. That is, each *S*'s score on each of the reading pretest trials was subtracted from his score on the comparable posttest trials. Thus if the learning on the naming trials did not influence the postreading task, the groups would fall at the zero level in Fig. 2.

The 30-sec. group manifested the greatest decrement attributable to the naming task; the 60-sec. group showed less decrement; and the 0-sec. group showed the least decrement. An analysis of variance was performed on transformed summed difference scores (Trials 1, 2, 3, 4, and 5 were combined for each group and treated as separate distributions). In order to meet the assumption of homogeneity of variance, the time scores resulting from combining trials were normalized by a square-root transformation. The resulting F was 5.97 ($p < .01$). It can be seen that the groups differ in the amount of interference manifested.

One striking phenomenon indicated in Table 1 and depicted in the left-hand portion of Fig. 2 is the performance of the 0-sec. group. This group showed relatively little interference on Trials 1 and 2. However, on Trial 3 some rise in interference is evidenced. Thereafter, interference increased only slightly on Trials 4 and 5 until the interference level slightly surpassed that of the 60-sec. group on Trial 5. By means of chi square, a test was made among the groups as to the point of maximum interference. This analysis was performed by counting the number of individuals in each group that showed maximum interference on the first trial as compared to all other trials. The resulting χ^2 was 6.67 which for 2 *df* falls between the .05 and .02 significance levels. Thus it seems that there is a significant shift in the point of maximum interference depending upon intertrial distribution in learning the interfering task.

TABLE 1

DIFFERENCES BETWEEN SCORES ON PRETEST AND POSTTEST TRIALS ON THE READING TASK

Intertrial Interval	Trials									
	1		2		3		4		5	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Experiment I										
0	10.66	10.57	11.00	16.11	17.20	22.84	19.60	21.23	20.73	9.97
30	64.80	58.54	50.20	40.28	46.07	15.89	39.40	24.85	28.74	16.13
60	38.73	19.95	24.60	22.14	25.80	26.08	21.87	20.98	19.60	18.35
Experiment II										
0	87.40	55.21	51.53	42.38	38.80	42.62	27.33	28.05	21.13	24.47
30	89.67	38.92	62.40	34.41	41.13	31.38	31.86	26.36	31.13	20.80
60	86.93	61.22	53.40	47.68	40.73	43.08	38.34	34.35	30.80	29.81

Table 1 and the curves shown in the right portion of Fig. 2 indicate the amount of interference manifested by the groups in Exp. II when a 24-hr. interval is introduced between the naming task and reading posttest. It it apparent that the 24-hr. rest interval tends to eliminate differences between the groups as to the amount of interference produced. An analysis of variance on transformed total scores (square root) confirmed this observation, i.e., the over-all difference was not significant. In addition, although the data are not shown here, it was found that even though 10 relearning trials on the reading task were given in Exp. II, interference was never completely eliminated.

Returning again to Fig. 2, it can be seen that all groups showed greater interference in Exp. II than in Exp. I. The only factor to account for this is the difference in the time interval between the two experiments.

Discussion

Both experiments demonstrated that it is possible to interfere with a well-established verbal habit. Both of the variables manipulated in this study affected the amount of interference produced. In Exp. I, where the reading posttest followed the naming task immediately, it was found that a 30-sec. distribution interval between trials on the naming task produced the most interference. In Exp. II, where there was a 24-hr. interval between naming and reading posttest trials, the degree of distribution on the naming task did not produce a differential interference effect on the reading task. The best explanation for these facts seems to be provided by a reactive inhibition hypothesis (2).

This inhibition (I_R) accumulated as *S*s performed on the naming task as evidenced by (*a*) decreasing performance levels within each day after Day 1 by all groups, (*b*) greatest performance decrement manifested by the 0-sec. group, and (*c*) increased performance level by each group at the start of each day. Thus it would seem that the habit strength of the naming response increased each day although temporarily depressed within each day by I_R. Now, when the reading posttest is introduced almost immediately (Exp. I), not all of the I_R developed on the naming task had dissipated. Thus, habit strength on this task was still depressed and interference with reading was minimal. That some inhibition remained from the naming task at the time of the reading task is especially apparent in Exp. I when the performance of the 0-sec. group is considered. In this particular case, interference was minimal on the first two reading posttest trials. It was not until Trial 3 was reached that any apparent interference was obtained in this group. By this time, this group had had a 1½-min. rest (30 sec. between naming and reading trials, 30 sec. between Trials 1 and 2, and 30 sec. between Trials 2 and 3 on the reading posttest) which allowed some of the inhibition to dissipate. Thus habit strength of the naming task appears to actually increase during this time producing more and more interference with the reading posttest.

Thus the full interfering effect of the naming response is not evidenced until measurement of the reading response is made after all I_R had dissipated. The lack of differences between the 0-, 30-, and 60-sec. groups in interference in the reading response in Exp. II is in agreement with this interpretation. The only fact that is not consistent with the I_R interpretation is that the posttest reading performance of the 30-sec. group for Exp. I is superior to that of the 60-sec. group. No explanation can be offered for this contradiction to I_R theory.

There is a possibility that the difference in the amount of interference manifested in Exp. I and Exp. II might be explained by I_R generalization. That is, the greater I_R in Exp. I, since it had not completely dissipated, tended to depress the reading response as well as the naming response. However, such an ex-

planation cannot account for the increase of interference in the 0-sec. group in Exp. I.

Summary

This study was concerned with the production of a decrement in the well-established reading habit. To produce this decrement, *Ss* were presented with a list of 144 color words which were printed in different colors. The *Ss* read the color words (reading pretest) for five trials and followed this with 60 trials (15 a day for four days) naming the color in which the color word was printed. Following this, *Ss* returned to the reading task (reading posttest). The intertrial distribution of the naming task and the amount of time between the naming task and the reading posttest were the independent variables.

The major results were as follows:

1. Distributed practice up to 60 sec. facilitated performance on the naming task.

2. When the reading posttest was introduced immediately (30 sec.) after the last trial on the naming task interference was the greatest for the 30-sec. group and least for the 0-sec. group. However, when the time interval between the naming and reading posttest trials was extended to 24 hr., interference was equal for all groups.

3. With the 30-sec. interval between reading trials, the point of maximum interference occurred earlier in the reading posttest trials if distributed practice had been given on the naming task.

4. Interference dissipated more rapidly when a 24-hr. interval rather than a 30-sec. was allowed between naming task and postreading task.

5. A reactive inhibition hypothesis was utilized to account for the obtained results.

References

1. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
2. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
3. Stroop, J. R. Studies of interference in serial verbal reactions. *J. exp. Psychol.*, 1935, **18**, 643–662.

(Received July 1, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

CLASSICAL AND INSTRUMENTAL EYELID CONDITIONING

GREGORY A. KIMBLE, LUCIE I. MANN, AND ROBERT H. DUFORT[1]

Duke University

Adherents (**10**) and opponents (**1**) to reinforcement theory have both argued that, according to the reinforcement point of view, classical conditioning in which the UCS_s is presented on every trial should be more effective in establishing a CR than an instrumental avoidance conditioning procedure in which the occurrence of a CR prevents the occurrence of the UCS. This inference from reinforcement theory is based on the fact that the classical procedure introduces reinforcement on every trial, whereas the instrumental procedure omits reinforcement whenever a CR occurs. The best evidence for this deduction is to be found in an experiment by Logan (**10**) who obtained better eyelid conditioning with a classical conditioning procedure than with an instrumental conditioning one. Such evidence, however, does not constitute *preferential* support for the reinforcement position as opposed, for example, to S-R contiguity theory. Sheffield (**11**), who adopts the latter position, also comes to the conclusion that conditioning should be faster under the classical procedure which assures the occurrence of a vigorous response on every trial. Thus, for different reasons, contiguity theory and reinforcement theory lead to the same prediction. To determine whether stimulus-response contiguity, reinforcement, or some other factor is the critical one requires a more detailed experimental analysis of behavior in the classical and instrumental conditioning situations. The four experiments reported in this paper are a step in that direction.

[1] Financial support for this research project was provided by the Duke University Research Council and the National Science Foundation.

Experiment I

The purpose of this experiment was to compare eyelid conditioning under classical and instrumental procedures. Since some such factor as delay of reinforcement might conceivably operate to make the more effective procedure at one interstimulus interval the less effective one at some other interval, the comparison was made for three different CS-UCS intervals.

Method

The apparatus, general procedure, and source of *S*s were identical for all four experiments. Hence, we shall give a general description here and supply necessary additional details for the later experiments when we come to them.

Apparatus.—The *S* sat at a table with his chin and forehead supported by an adjustable headrest, looking into a booth which (*a*) kept *E* and the timing apparatus out of *S*'s view, (*b*) provided fairly constant conditions of general illumination, and (*c*) provided a background against which the CS appeared. The CS was a dim (about 1.7 mL.) light of bluish color provided by an argon lamp. It was 1⅛ in. in diameter, 16 in. directly ahead of *S* and diffused by a ¼-in. ground glass screen which constituted the back wall of the booth. The UCS was an air puff delivered to the right eye through a $\frac{3}{64}$-in. nozzle. It was magnetically released and provided by the fall of a 90-mm. column of mercury in a manometer. The durations of the CS (250 msec.), interstimulus interval, and intertrial interval (15 sec.) were controlled by a constantly running motor which activated a series of switches.

Without changing any other conditions, the apparatus provided an instrumental and a

classical version of the eyelid conditioning experiment. In the instrumental condition the UCS followed the CS only if *S* failed to make a response to the CS. In the classical condition the UCS followed the CS regardless of *S*'s behavior.

A very light false eyelash made of aluminum foil attached to *S*'s right eyelid interrupted the focal point of a light beam directed at the element of a photocell. The beam was focused on the eyelash about 1¼ in. from the eye. When *S* blinked approximately .075 in., the light reached the photocell and activated a circuit controlling the recording mechanism and, in the instrumental condition, preventing the delivery of the air puff.

The occurrence of a blink within the CS-UCS interval was signalled by a light on a panel of lights. Because of the recording system used, the records were of the "all or nothing" variety, providing measures neither of magnitude nor latency. Of these the latter is potentially the more serious because of the possibility of counting as CR's certain UCR's to light described by Grant and Norris (**6**) or voluntary responses of the sort described by Spence and Taylor (**12**). Although our data probably contain some responses which are not true CR's, it seems unlikely that they are very numerous. The number of UCR's to light presumably was minimized for the following two reasons: (*a*) the *S*s were run in the light, and (*b*) the intensity of the CS was low. The possibility of a contamination of the data by voluntary responses remains. Again it seems unlikely that these responses were very important for the following reasons: (*a*) According to Spence and Taylor, voluntary responders are individuals who are attempting to cooperate with *E* because of a knowledge of eyelid conditioning gained from a classroom discussion of conditioning. Except for a very few *S*s derived from sources other than the introductory course, the *S*s in these studies were run before any classroom discussion of conditioning; (*b*) These voluntary responders make up only about 10% of the population and thus cannot be held responsible for all of the effects obtained in these experiments. In any event, the number of responses which are not true CR's must have been randomly distributed among the various groups and could not account for the differences obtained. Finally, (*c*) the good agreement of the data obtained in these experiments with other data which were presumably uncontaminated by these particular artifacts (e.g., **8, 10, 13**) indicates that we are dealing with CR's rather than some variety of UCR.

Subjects and design.—The *S*s were 36 undergraduate students at Duke University. Eighteen were conditioned by the classical procedure and 18 by the instrumental procedure. Each major group was subdivided and 6 *S*s, 3 men and 3 women, were conditioned at 3 different CS-UCS intervals (500, 800, and 1500 msec.).

Procedure.—Instructions to *S*s were the following: "This experiment is concerned with the reaction of the human eye to different kinds of stimulation. The stimulation will be provided by a blue light and an air puff. The light will appear in the center of the window directly in front of you, every so often. The air puff will be delivered to your right eye at various times during the experiment. The reaction that I want to measure is the eyeblink. To do so it is necessary to attach a false eyelash to your right eye (*E* attaches eyelash).

"In this experiment, your blinking is recorded by an electric eye. For this reason, it is important to fix the position of your head quite exactly. To do this we use the head and chin rest. During the experiment you are to sit here with your chin in the chin rest and your head pressing against the upper part of the head rest. You are always to look at the place in front of you where the light appears (*E* adjusts *S* to apparatus).

"The most important thing for you to do is to keep comfortable and relaxed. For instance, when you are looking at the light, do not inhibit your blinking. If you feel your eyes closing or about to close, do nothing to prevent it. On the other hand, do not blink purposely to catch the air puff on the outer part of your eye. You do not have to think about the experiment while you serve as subject. In fact, you will be of more help if you think about something else."

After these instructions had been presented and *S* had been adjusted to the apparatus, the CS was presented alone ten times. The four *S*s who gave six or more responses to the CS on these trials received ten additional adaptation trials which reduced the number of such responses to four or fewer. Following a 1-min. rest, *S* received 60 training trials with an intertrial interval of 15 sec. except for a 1-min. rest after Trials 20 and 45. A series of 30 extinction trials began 15 sec. following the last training trial. No special rest occurred between training and extinction.

Results[2]

The data most directly pertinent to the problem under investigation appear in Fig. 1. In that graph the

[2] Detailed tables of the means and variances for Exp. 1, and for Exp. 2, 3, and 4, have been deposited with the American Documentation Institute. Order Document No. 4516, remitting $1.25 for microfilm or $1.25 for photocopies.

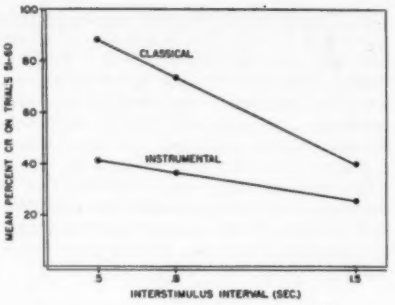

FIG. 1. Mean percentage of conditioned responses on Trials 51–60.

points are mean percentages of CR's obtained on the last ten conditioning trials, where the learning curves for the various groups were at or near an upper limit. The important features of the two curves are these: (*a*) The amount of conditioning at every point is greater for the classical condition than for the instrumental condition. (*b*) The amount of conditioning decreases with increases in the interstimulus time. Analysis of variance supports these interpretations. The *F* for the effect of interstimulus interval was 5.70 which for 2 and 30 *df* is significant at the .01 level. The *F* for experimental condition was 16.92 which for 1 and 30 *df* is significant at the .001 level. The interaction term does not approach significance ($F = 1.59$), indicating that the two functions are parallel within the limits of sampling error.

In his investigation of this problem, Sheffield (**11**) proposed that the instrumental version of conditioning should be one in which *S* goes through a series of alternating conditioning and extinction phases. His reasoning was that, on trials where the CS and UCS occur together, conditioning takes place. Trials on which *S* avoids the UCS should lead to extinction. Detailed examination of the records from *S*s conditioned by the instrumental procedure seems to support his analysis. A typical record began with few CR's and passed through a point where four or five CR's were alternated with four or five failures. Toward the end of the record, these runs became shorter and one or two CR's were alternated with one or two failures. The classical groups, by contrast, tended to show increasingly longer runs of CR's. Using a runs test described by Dixon and Massey (**3**), it is possible to evaluate such differences statistically. This test begins by assigning a Z-score to the number of runs in each *S*'s record in a way which takes the number of CR's into account. These *Z*-scores then may be treated as ordinary scores and used to provide an unbiased test of the number of runs in any consecutive set of data. In this experiment, the instrumental groups' data contained significantly more runs than the data for the classical groups. A *t* test applied to these *Z*-scores produces a *t* of 2.80 ($p = .01$). It is particularly important here to mention that this does *not* simply prove that the classical groups learned more than the instrumental groups, since the computation of the *Z*-score eliminates differences in number of CR's. This test, instead, reflects a different *pattern* of performance under the two reinforcement conditions.

Discussion

These results confirm those of Logan (**10**), and contrast sharply with those of such other investigators as Brogden, Lipman, and Culler (**2**), Gibson (**4**), Hunter (**9**), and Sheffield (**11**) who obtained faster conditioning by the instrumental procedure. Just why such conflicting results should have come from these different experiments cannot be determined on the basis of currently available data.

Although the present results are compatible with predictions from reinforce-

ment theory, they are equally in line with predictions from contiguity theory. Whether the differences in conditioning obtained in the classical and instrumental groups in this experiment resulted from different amounts of reinforcement, from different frequencies of occurrence of the blink in close temporal contiguity with the CS, or from some other factor, cannot be determined from the data. The three experiments which follow attempt to determine the nature of the process involved.

Experiment II

Previous investigators of the relative effectiveness of classical and instrumental conditioning have never raised what might be regarded as a basic question: Is the difference in behavior produced by these two procedures a difference in learning, performance, or both? This experiment is an attempt to answer this question. It employed a factorial design in which different groups of *Ss* began conditioning under the classical or instrumental procedures. Midway in the experiment, these two groups were split, half of each group continuing under the original condition, half switching to the alternate condition. In the statistical analysis of the results of such experiments, the important data are those obtained in the second series of trials. Differences which reflect a significant effect of the conditions existing prior to the shift are taken to indicate differences in learning; differences attributable to conditions existing after the shift are taken to reflect differences in performance.

Method

The apparatus was that used in Exp. I. The *Ss* were 24 Duke University undergraduates who served under four conditions. All *Ss* received ten adaptaion trials with the CS alone. One *S*, who responded six times, received an additional ten adaptation trials. Following these trials, all *Ss* received 120 conditioning trials with a CS-UCS interval of 500 msec. and an intertrial interval of 15 sec. The *S* was allowed to rest for 1 min. after Trials 20, 45, and 80. Twelve *Ss* received their first 60 trials under the classical condition; the other 12 began in the instrumental condition. Beginning with Trial 61, half of the classical group continued on the classical condition; the other half changed to the instrumental condition. Beginning with Trial 61 the instrumental group was similarly split. We shall designate the four groups thus generated as *instrumental-instrumental*, *instrumental-classical*, *classical-classical*, and *classical-instrumental*.

Results and Discussion

Learning curves for the four groups appear in Fig. 2. These curves show that: (*a*) The difference between classical and instrumental conditioning obtained in Exp. I reappears here. (*b*) With the change from the instrumental to the classical condition there is a marked and sudden increase in the level of performance. (*c*) The shift from classical to instrumental conditioning produces a much less dramatic decrease in performance. For Trials 61–120, the difference between the mean number of CR's for the classical-instrumental group and the classical-classical group, though small, is significant ($t = 8.08$; $p = .001$).

Again the runs test described in con-

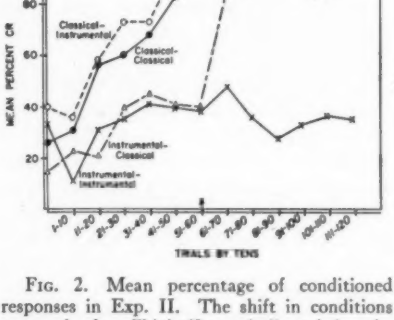

Fig. 2. Mean percentage of conditioned responses in Exp. II. The shift in conditions occurred after Trial 60, as indicated by the arrow. Points at the extreme left are for test trials.

nection with Exp. I reveals a greater tendency toward cycles of CR's and failures to respond in the instrumental groups' behavior on Trials 1–60 than for the classical groups on the same trials ($t = 2.40$; $p = .025$). It would be interesting to know whether this tendency holds up for the classical-instrumental group in Trials 61–120, but, because of the large number of CR's on these trials, the test is inapplicable.

An analysis of variance performed on the total number of CR's occurring on Trials 61–120 produced a highly significant ($F = 250.13$; $p = .001$) interaction between experimental condition on Trials 1–60 and 61–120, necessitating the use of the interaction variance as the error term to evaluate the main effects. Such a test shows that neither the experimental condition on Trials 1–60 or 61–120 produced a significant effect. The two F's were 1.09 and 1.40. Thus the statistics appropriate to the experimental design fail to indicate whether the influence of the classical and instrumental conditioning procedures is on learning or on performance. One feature of the conditioning curves suggests, however, that the results may be explained entirely in terms of performance variables, probably motivational ones. The reference here is to the rapid rise in the curve for the instrumental-classical group when these *S*s shifted from one conditioning procedure to the other. If we were dealing with changing habit strengths one would predict, at least on the basis of Hull's (7) theory, that the increase would have been much slower. Of the main classes of variables influencing performance, motivational factors seem the more likely.

It is also clear that, whatever the factor responsible for the results of these experiments, it develops with practice. Such a consideration leads to the suggestion that we are dealing with an acquired drive based on the air puff, which reaches different levels under our two major experimental conditions. It is at this point that the cycles in the behavior of *S*s in the instrumental condition become critical. What we suppose to be happening in that condition is this: On trials where *S* fails to make a CR he receives an air puff and there develops an increment of acquired motivation. On trials where *S* makes a CR this acquired drive undergoes some degree of extinction because the air puff is not delivered. When the extinction process decreases drive to the point where *S* again fails to respond, the UCS appears again, and the acquired drive begins to be reinstated. Successive conditioning and extinction of a learned motivation, thus, keep its average value low in the instrumental conditioning procedure. It seems likely that the only sense in which the two experimental procedures influence learning is in their effect on an acquired drive. With that exception the argument above is the same as that offered by Sheffield (11) to account for certain results of his experiment.

This analysis will handle the results of this experiment in the following way. When *S*s in the instrumental-classical condition enter the series of trials of classical conditioning, the extinction process no longer occurs and, as drive increases, so does performance up to the level of the classical-classical group. When *S*s in the classical-instrumental group enter the series of instrumental trials the successive conditioning and extinction process begins. Since the earlier classical trials have already brought drive to a high level, however, the extinction and conditioning processes function around a higher value, with the result that there is some decrease in performance, but not so much as the increase when the shift is from instrumental to classical conditioning.

This explanation for the effects occurring in these experiments relies heavily upon two assumptions: (*a*) that the cycles of one or two reinforced and nonreinforced trials have a significant effect, and (*b*) that this effect is upon some

motivational process which can vary independently of a more basic learning process. Experiments III and IV test these assumptions.

Experiment III

The purpose of this experiment was to introduce, *as an independent variable*, the patterns of reinforcement and nonreinforcement observed in the response records of *S*s conditioned by the instrumental procedure. The method employed was to condition a group of *S*s on a partial-reinforcement schedule which omitted air puffs at the points where *S*s in the previous instrumental conditioning groups had avoided them. A second group was conditioned on a different partial reinforcement schedule which omitted the same number of reinforcements at points where the *S*s in the previous groups had received them.

Method

The apparatus was the same as that used in Exp. I and II. The *S*s were 12 Duke University undergraduates who served under two different partial reinforcement versions of the classical conditioning situation. One additional *S* who responded 5 times in the adaptation series was eliminated. Each *S* received 60 training trials and 30 extinction trials.

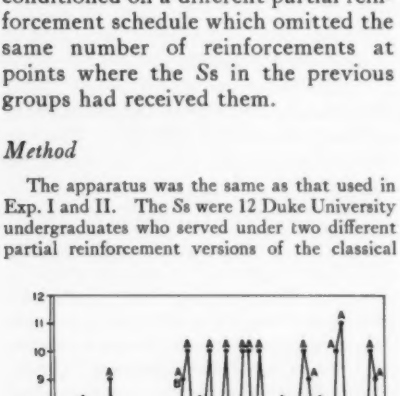

Fig. 3. Curve of total number of conditioned responses produced on each trial by 18 *S*s in the 500-msec. instrumental conditioning procedure in Exp. I and II. Points labeled "A" and "B" indicate trials on which the air puff was omitted in two partial reinforcement schedules.

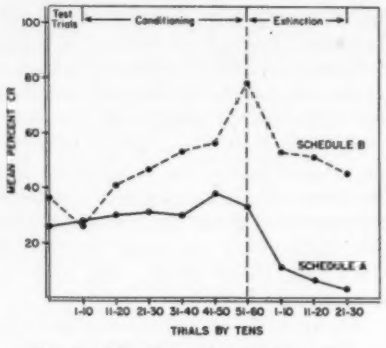

Fig. 4. Learning and extinction curves for Exp. III.

The procedure for determining the two partial reinforcement schedules is illustrated in Fig. 3, which presents the total number of CR's obtained on each trial from the 18 *S*s in the instrumental conditioning groups with a 500-msec. CS-UCS interval in Exp. I and II. The points marked "A" indicate the trials on which the air puff was omitted for one group. These trials are those on which the air puff was most frequently avoided in each block of 10 trials by the earlier *S*s. The points marked "B" indicate the trials on which the air puff was omitted for the second group of *S*s. These tend to be the trials on which the air puff was most frequently received by the earlier *S*s. In constructing these schedules the two criteria were (*a*) that every block of 10 trials should omit about the same number of air puffs as had been avoided by the average *S* in the previous experiments, and (*b*) that neither schedule should ever omit more than two air puffs in succession. It should be noted that Schedule A omits two successive air puffs on seven occasions. In Schedule B, this happens only twice.

Results

Learning and extinction data for these two groups are presented graphically in Fig. 4. Obviously, the two different schedules produced quite different courses of acquisition and differential resistance to extinction. Using the mean total number of CR's in acquisition and extinction as measures of learning, and testing for the

significance of the differences between these means, we obtain the following results: For conditioning, $t = 2.96$, $p = .02$; for extinction, $t = 7.42$, $p = .001$.

The assumption behind the use of the two partial reinforcement schedules was that Schedule A would omit air puffs at the points where the same *S*s might have avoided them if they had been conditioned by the instrumental procedure, and that Schedule B would systematically introduce air puffs on these trials. To determine whether the schedules had actually accomplished this, we compared the percentage of the total CR's produced by each group on the 21 trials on which the air puff was omitted. Of the relatively smaller number of CR's produced under Schedule A, 38% occurred on these critical trials. Of the relatively larger number of CR's produced under Schedule B, only 26% occurred on these trials. The difference between these mean percentages is significant at the .05 level ($t = 2.37$). In computing this statistic the percentages of responses on the critical trials for each *S* were taken as scores. During the 60-trial conditioning series, the data for the group conditioned under Schedule A correspond fairly well to those obtained from the instrumental group with an interstimulus time of 500 msec. in Exp. I. The latter group produced slightly more CR's, but the difference is not significant ($t = .87$; $p = .50$). There is, however, a significant superiority of the instrumental group in Exp. I in extinction ($t = 2.27$; $p = .05$).

On the assumption that Schedule B in this experiment probably resembled the partial reinforcement schedule used by Humphreys (**8**) in his study of 50% and 100% reinforcement in eyelid conditioning, we compared the results obtained under this schedule with those produced by the classical group with a 500-msec. interstimulus interval in Exp. I. This comparison indicates that the present results are substantially the same as Humphreys'. The group from Exp. I produced a slightly greater number of CR's in the training trials, but the difference was not significant ($t = 1.37$; $p = .20$). The group from Exp. I, however, was less resistant to extinction ($t = 2.99$; $p = .02$). To make this test, the scores used were difference scores between number of CR's in training and extinction. Using this measure extracts the correlation between conditioning and extinction measures from the error term used in the t test and is the approximate equivalent of an analysis of covariance.

Discussion

The outcome of this experiment shows that the reduction in the level of conditioning by the instrumental procedure must be assigned to the precise pattern of omitted air puffs. It also shows, incidentally, that the *number* of paired presentations of CS and UCS cannot be the important variable since the number was the same in the two groups. Since one difference between the two schedules was that Schedule A, which produced poorer conditioning, tended to omit more pairs of air puffs than Schedule B, it is possible that as few as two omitted air puffs or extinction trials may produce a significant effect. What remains to be demonstrated is that the effect is specifically on some motivational process.

Experiment IV

This experiment is a test of the assumption that performance in the eyelid conditioning situation depends jointly upon an associative factor and an acquired drive, and that these two processes are relatively independent of each other. The procedure involved an attempt to increase the acquired drive without altering the

level of conditioning. The method of doing this was to present the UCS without the CS on several trials after a short series of preliminary conditioning trials.

Method

Three groups of seven Duke University undergraduates were conditioned by the classical procedure with a 500-msec. interstimulus interval. Each group received 60 trials at 15-sec. intertrial intervals, followed by 30 extinction trials. One group was a control group which received 60 conditioning trials in which the CS and UCS were paired on every trial. The other two groups received 20 such trials. At this point the apparatus was turned off to eliminate auditory cues and one group (the 10-trial group) received 10 manually delivered presentations of the UCS alone at approximately 15 sec. intervals; the other group (the 20-trial group) received 20 presentations of the UCS alone. The *Ss* were told that these were administered to "check the recording apparatus." Following these presentations of the air puff alone, each group received enough paired presentations of the CS and UCS together to bring the total number of trials to 60 for each group. The 30-trial extinction series then followed immediately.

This procedure is almost exactly the same in the experimental conditions as that used by Grant (**5**) to obtain pseudo-conditioning, and raises what is potentially an extremely embarrassing question. As a control, Grant ran a group of *Ss* which merely stayed in the apparatus for the length of time required to present the unpaired air puff stimulations to the experimental group. He found that this group showed an increment in number of "CR's" very similar to that obtained from the experimental group. In short, in his experiment, the presentation of the air puff had nothing to do with the increase in number of blink responses. Subsequent research (**6**) has identified the responses producing these increases as beta responses, sensitized by dark adaptation. It has previously been pointed out that the procedure used in these experiments was not likely to elicit many of these responses because *Ss* were not dark adapted and the CS intensity was low. As a further check, the present experiment employed two additional control groups designed to determine whether any effects obtained depended on the use of the air puff. One of these control groups corresponded to that used by Grant (**5**). Six *Ss* received 20 conditioning trials and then sat in the apparatus looking at a fixation point for the amount of time required to present the UCS 20 times to a 20-trial group. The other group of six *Ss* was subjected to exactly the same procedures as were used with the 20-trial group, except that the intensity of the air puff during Trials 21–40 was of 45 mm. rather than 90 mm. These groups provide the means of determining whether the presentation of the air puff is the critical factor in producing whatever results come from this experiment. If the presentation of the air puff is critical, we anticipate: (*a*) little or no increase in number of CR's for the no-puff, fixation group, and (*b*) a greater effect in the 90-mm., 20-trial group than in the 45-mm. group.

Results

The basic data obtained in the experiment appear in Fig. 5. The curve which proceeds without interruption is that for the control group which received 60 pairings of CS and UCS. The curves for the other two groups are broken between Trial 20 and the point at which the CS was reintroduced. The striking feature of these curves is that there is no evidence, either in acquisition or extinction, that the omission of the CS on a series of trials has any detrimental effect on learning. If anything, the opposite is true. The superiority of

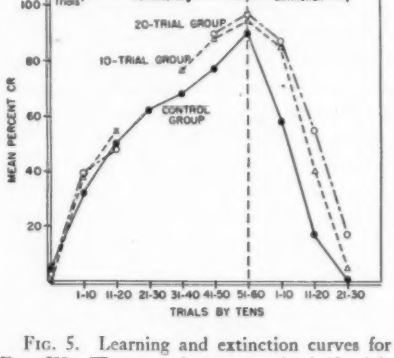

FIG. 5. Learning and extinction curves for Exp. IV. The control group received 60 trials in which CS and UCS were paired. The CS was omitted on Trials 21–30 for the 10-trial group and on 21–40 for the 20-trial group.

the two experimental groups during the last part of training is not statistically significant, but the superiority of these groups during extinction is significant ($F = 7.58$; $p = .005$). The most likely explanation for this appears to be that *Ss* in the experimental groups receive the air puff on the open eye more often than *Ss* in the control group. This would probably be the same as increasing the intensity of the air puff.

There is, as the curves in Fig. 5 suggest, a significantly greater increase in performance for the 20-trial group than for the 10-trial group from Trials 11–20 to the point where the CS was reintroduced. The mean gain in number of CR's is 2.54 for the 10-trial group and 4.00 for the 20-trial group ($t = 4.54$; $p = .001$). This result apparently demonstrates the acquisition of an acquired drive with practice.

The results of the 20-trial group are re-plotted in Fig. 6, together with the data for the two additional control groups. These curves provide clear-cut evidence for the two effects demanded by the assumption that the results just presented depend upon the unpaired air-puff stimulations: (*a*) The increase in number of CR's for the no-puff control group is slight (5.0%) and only approaches statistical significance ($t = 2.27$; $p = .10$). (*b*) There is a greater increase in percentage of CR's for the 90-mm. puff group than for the 45-mm. puff group. The mean increase for the first of these groups was 40.0%, that for the second was 21.4% ($t = 2.78$, $p = .02$). These findings seem to rule out the possibility that these data reflect only the sensitization of some variety of UCR to light. If additional evidence is needed, the extinction data provide it. Grant's (**5**) results for groups similar to these show no evidence whatsoever for extinction

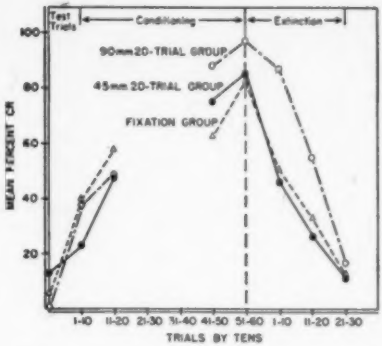

FIG. 6. Learning and extinction curves for the 20-trial group in Exp. IV together with data for the two control groups which received a 45-mm. air puff or no air puff on Trials 21–40

of the sensitized beta response. The fact that the extinction curves in this experiment are fairly typical indicates that the responses are true CR's.

Discussion

The relevance of the results of this experiment to the previous ones in this series lies chiefly in their support for the existence of two, independently manipulable, processes in eyelid conditioning. The first is an *associative* factor which depends upon the presentation of the CS and UCS together. Apparently it develops rapidly. Certainly, after 20 conditioning trials, whether the CS is presented or not is irrelevant so long as the UCS is presented. The behavioral changes which are usually plotted as learning curves for eyelid conditioning probably reflect mainly the development of a second, *performance*, factor which is under the control of the UCS. This last fact supports the interpretation that the performance factor is a motivational one.

GENERAL COMMENT

Previous discussions of the probable importance of reinforcement in defense conditioning by theorists sympathetic to the reinforcement point of view have begun with the assumption that the

termination of the UCS provides a potential source of drive reduction and, therefore, of reward. They have then proceeded to the prediction that the classical version of such conditioning should lead to more rapid *learning* than the instrumental version because it provides for a greater number of reinforcements. The results of these experiments call this explanation into question in two ways. First, they indicate that the effect of the two experimental conditions is not primarily upon learning, but upon performance. Second, they suggest that, beyond the first few pairings of CS and UCS, the number of reinforcements has no significant effect on learning.

With respect to the details of the usual reinforcement explanation, then, the results of these experiments are substantially negative. This is not, however, to say that reinforcement plays no role in the establishment of classically conditioned responses. There remains the possibility that reinforcement is involved in the development of the acquired drive assumed to be operating in these experiments. Whether such function of reinforcement is likely is an empirical question which can be answered by further research.

Summary

The experiments in this series compared the effectiveness of classical and instrumental procedures for conditioning the eyeblink response of human *Ss*. In Exp. I, 36 *Ss* were conditioned by these two procedures with 500-msec., 800-msec., and 1500-msec. interstimulus intervals. Results showed (*a*) poorer conditioning under the instrumental condition ($p = .001$), (*b*) the usual decrease in amount of conditioning with increasing interstimulus intervals for both conditions ($p = .01$), and (*c*) a tendency toward cyclical alternations of runs of CR's and failures to respond in the instrumental condition only.

In Exp. II, 24 *Ss* were used with half of them beginning on the instrumental condition, half on the classical condition. Midway in conditioning, half of each group was switched to the other condition. Results show a large *increase* in performance when the instrumental group is changed to the classical condition ($p = .001$) and a smaller *reduction* in performahce when the classical group is changed to the instrumental condition ($p = .001$). Again, there is evidence for cyclical alternations in the performance of the instrumental groups.

In Exp. III, 12 *Ss* were conditioned with two different partial reinforcement schedules where the air puff was otherwise unavoidable. One of these schedules attempted to omit air puffs where the previous instrumental groups had avoided them; the other attempted to introduce air puffs at these points. Both conditioning ($p = .02$) and extinction ($p = .001$) were superior under the latter schedule.

In Exp. IV, 21 *Ss* were conditioned under the classical conditioning procedure. In one group the CS and UCS were presented together on every trial. In two other groups, the CS was omitted on Trials 21–30 and 21–40, and the UCS was presented alone 10 or 20 times in an effort to manipulate *S*'s motivation without changing the amount of conditioning. The results of this experiment demonstrated that the omission of the CS did not interfere with the progress of conditioning. Data from two additional control groups show that these effects depend upon the presentation of the UCS and not upon an artifactual sensitization of a UCR to light.

The interpretation offered for these results is that the instrumental conditioning procedure produces cycles of conditioning and extinction of an acquired drive, which result in lowered performance although the amount of learning is equal to that in the classical procedure.

References

1. Birch, H. G., & Bitterman, M. E. Reinforcement and learning: the process of sensory integration. *Psychol. Rev.*, 1949, **56**, 292–308.
2. Brogden, W. J., Lipman, E. A., & Culler, E. The role of incentive in conditioning and extinction. *Amer. J. Psychol.*, 1938, **51**, 109–117.
3. Dixon, W. J., & Massey, F. J., Jr. *Introduction to statistical analysis*. New York: McGraw-Hill, 1951.
4. Gibson, E. J. The role of shock in reinforcement. *J. comp. physiol. Psychol.*, 1952, **45**, 18–30.
5. Grant, D. A. The pseudo-conditioned eyelid response. *J. exp. Psychol.*, 1943, **32**, 139–149.
6. Grant, D. A., & Norris, E. B. Dark adaptation as a factor in the sensitization of the beta response of the eyelid to light. *J. exp. Psychol.*, 1946, **36**, 390–397.

7. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
8. Humphreys, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25,** 141–158.
9. Hunter, W. S. Conditioning and extinction in the rat. *Brit. J. Psychol.*, 1935, **26,** 135–148.
10. Logan, F. A. A comparison of avoidance and nonavoidance eyelid conditioning. *J. exp. Psychol.*, 1951, **42,** 391–393.
11. Sheffield, F. D. Avoidance training and the contiguity principle. *J. comp. physiol. Psychol.*, 1948, **41,** 165–177.
12. Spence, K. W., & Taylor, J. Anxiety and the strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42,** 183–188.
13. White, C. T., & Schlosberg, H. Degree of conditioning of the GSR as a function of the period of delay. *J. exp. Psychol.*, 1952, **43,** 357–62.

(Received August 13, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

AN EVALUATION OF THE EFFECT OF INDUCED TENSION ON PERFORMANCE

LYLE E. BOURNE, JR.[1]

University of Wisconsin

In the past, many studies (e.g., **1, 3, 4, 6**) have been directed toward the determination of the effect of experimentally induced tension on performance. The results of these investigations have not been unequivocal, some experimenters reporting facilitation, some inhibition, and some no appreciable effect. The outcome of this diversity in data has been to focus attention on "the problem of defining the conditions which modify the influence of tension . . ." (**3,** p. 347).

Experimentation directed along this line has left one very important problem to be clarified. The assumption in most of the previous work has been that tension exerts some direct influence on acquisition of skill. The general conclusion has been that under specified, optimal conditions, tension will facilitate *learning*. Recently, Meyer hypothesized that induced muscular tension facilitates *responding* "by making the motor system available to an input, not by operating upon the input" (**9,** p. 209). The effect of tension is not upon an individual's ability to learn but rather upon his ability to respond. The theory can account for changes in magnitude and latency of a response but not for changes in rate of acquisition.

Data not in agreement with this hypothesis are presented in a study by Bills (**1**), who found paired-associate learning to be improved if *S* squeezed a dynamometer. Bills' conclusion was that induced muscular tension can affect directly the formation of habits. In the experiment, lists of paired associates were presented for a single trial under conditions of normal and induced tension. The "tension" group recalled the words significantly better than did the "normal" group. However, the recall trial followed almost immediately after the learning trial. Meyer doubts that induced tension was completely dissipated at the time of recall. If it were not, then the performance of the "tension" group during recall was influenced by this indissipated tension.

The present experimental design introduces time intervals of varied length between learning and recall. If Bills is correct, regardless of the time between learning and recall, *Ss* who learn under tension should recall better than those who learn under no tension. On the other hand, if tension facilitates only responding as proposed by Meyer, learning under tension should result in no better recall performance, provided enough time is allowed for induced tension to dissipate.

Procedure

Subjects.—Forty volunteer undergraduates in an introductory and an experimental psychology course served as *Ss*. These *Ss* were divided into four groups of 10 *Ss* each. The distribution of males and females was nearly equal in each group. The *Ss* served individually and, in general, for six successive days.

Apparatus.—A Hull-type memory drum, presenting each word for a 2-sec. duration, was

[1] The author wishes to express his gratitude to Dr. D. R. Meyer of the Ohio State University for direction of this research and theoretical implications of the data, and to Dr. E. J. Archer for assistance in design.

used. A list of ten paired adjectives and a series of 20 colored squares were placed on the same vellum tape. Either of these could be blocked from *S*'s view by means of a movable shutter which cut the memory-drum aperture to half its normal size. Thus, for an immediate shift from one activity to the other, *E* merely had to move the shutter from one side of the aperture to the other. The adjectives used were chosen from Haagen's list (**7**) and were of medium similarity value (2.0 to 3.0).

A Stoelting Co. hand dynamometer was mounted in a stationary position slightly to the right of the memory drum. A simple arrangement insured stable exertion by *S*. The pointer on the dynamometer was constructed of Plexiglas. At the tip of the pointer was an electrical contact. Over the face of the dynamometer, rotating on the center pole, was a circular Plexiglas sheet. At the periphery of the sheet was another electrical contact. This sheet could be fixed in any position, thus its contact could be placed at the value of ⅓ maximum tension for any *S*. When *S* squeezed to this value, the two contacts touched and closed an electrical circuit lighting an incandescent bulb. In this way both *E* and *S* were signalled, without watching the dynamometer scale, when *S* was at the correct tension level.

All sessions took place in an isolated, soundproof, air-conditioned room. A masonite partition separated *S* and *E* for the elimination of irrelevant stimuli and the standardization of presentation.

Task.—The *S*'s task was to learn paired adjectives. During an experimental session, *S* would be presented with a list of ten paired adjectives once. After a specified time interval, the list would be presented a second time during which *S* would try to recall as many of the associates as possible. Four groups of *S*s which differ in temporal condition of tension induction, were used in the experiment. Group T-T learned and recalled under tension, Group T-N learned under tension and recalled under no tension, Group N-T learned under no tension and recalled under tension, and Group N-N learned and recalled under no tension. Five time intervals between the two presentations of the list were used, 0, 30, 60, 120, and 240 sec. The immediate (0 sec.) condition duplicated the Bills design while the four other intervals were included to provide for the achievement of different levels of relaxation in the *S*s of Group T-N.

During the rest intervals, the time between the two presentations of the list, all *S*s were engaged in a color-viewing activity. In this task *S* watched a series of 20 colored squares, presented one at a time on the memory drum, without naming them. He was instructed at the outset of the experiment that he would be quizzed on each color series at the end of each session. The questions asked were always the same but the color series differed from day to day. In the typical color-naming activity, *S* is required to call out the names of the colors as they appear. Instead of an intervening task, this might better be labeled an interfering task. It is an activity which necessitates a response similar in many respects to the one required by the verbal learning activity. The color-viewing activity used in the present experiment did not require this overt response. It is submitted that the present color-viewing activity was an adequate *intervening* task.

Tension was induced by having *S* grip the hand dynamometer. The *E* determined each *S*'s maximum squeezing power during a practice session. During the experimental session, *S* was required to squeeze to ⅓ this maximum value. One-third maximum was chosen because it was found by Courts (**4**) to lie in the optimal tension range for verbal memorizing activity.

Experimental design.—The experimental design was of the greco-latin square type, rows corresponding to *S*s, columns to days, latin letters to time intervals between learning and recall, greek letters to lists of adjectives. Two 5×5 squares were used for a group of ten *S*s and these were replicated in all four groups. On the first of the six days that *S* was required to serve he was oriented to the apparatus and procedure, his maximum grpping power was determined, and he was given some practice in learning paired associates. On the second day he was presented with the first of the experimental lists.

Results

The results of the present study are shown in Fig. 1. The values plotted represent the mean number of adjectives correctly recalled by the ten *S*s within a group. It is apparent from this figure that a clear substantiation of the Meyer prediction was obtained. If *S* is recalling under tension (N-T, T-T) his performance is facilitated, as compared with recalling under no tension (N-N). This discrepancy remains approximately constant (.5 syllables) throughout the range of time intervals between learning and recall used in this study. The

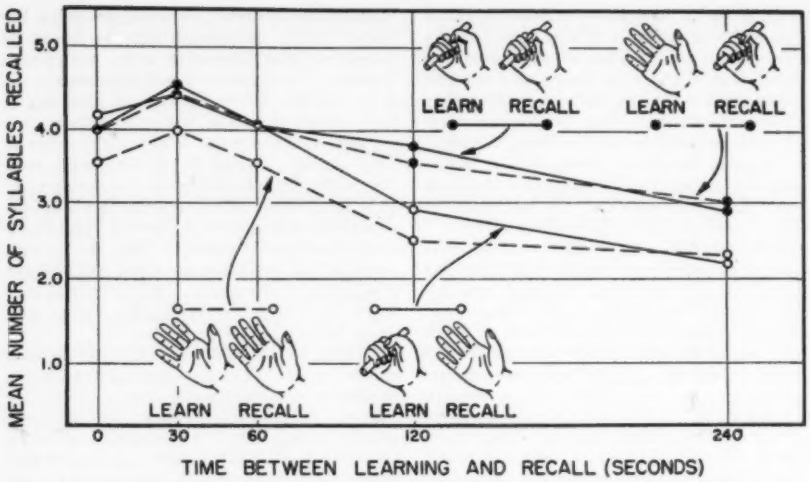

FIG. 1. Mean number of adjectives recalled as a function of the time between learning and recall and the temporal condition of tension induction. Every plotted point represents data from ten *Ss*.

Ss in Group T-N begin to relax after the learning trial. As the interval between learning and recall is increased, more of the induced tension is dissipated. The assumption is that *Ss* do not relax at once upon releasing the dynamometer. Thus, in the shorter interval conditions, *S* is recalling under some tension. The data show that as more time is allowed for tension dissipation, *S*'s performance is less facilitated, until, after a 4-min. interval, his recall score is no better than the score of an *S* who learned under no tension.

No attempt was made to match the four groups on ability to learn paired associates. However, a check on group equivalence was made by comparing mean number of trials needed by the *Ss* in each group to learn a practice list during the orientation on Day 1. The *Ss* in Group T-T required 7.4 trials, in Group T-N 7.0 trials, in Group N-T 7.1 trials, and in Group N-N 7.5 trials to learn this list to a criterion of one perfect repetition. The practice list of paired adjectives was presented in the same order on each trial which probably accounts for the rapid rate of learning.

An analysis of variance was performed on the recall scores. A summary of the analysis is presented in Table 1. Using the Pooled Residual

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE FOR RECALL SCORES.

Source	*df*	Mean Square	*F*
Groups	3	5.537	1.847
Ss/Groups	36	2.998	3.078**
Days	4	1.158	1.189
Time intervals	4	19.195	19.708**
List	4	8.158	8.376**
Day × Groups	12	1.868	1.918*
Time Intervals × Groups	12	0.422	0.433
Lists × Groups	12	1.118	1.148
Pooled residual	112	0.974	
Total	199		

* Significant at the .05 level.
** Significant at the .01 level.

as the error term for the mean squares of the correlated measures, it was found that the Day × Groups interaction is significant at the .05 level. The day-to-day variability is not significant ($F = 1.189$, 4 and 112 *df*) and there is no consistent trend in the scores as a function of practice. Since individual differences are significant ($F = 3.078$, 36 and 112 *df*, $p < .01$), the *Ss*/Groups mean square is the correct error term for testing group differences. It is not unexpected that the over-all difference between groups should turn out to be insignificant. According to Meyer, Groups T-T and N-T should yield essentially identical results. For a critical evaluation of Meyer's hypothesis, the difference between Groups T-N and N-N at the 0- and the 240-sec. intervals and the difference between pooled Groups N-T and T-T and pooled Groups T-N and N-N at the 240-sec. interval should be tested. Separate *t* tests were used at these points. The difference between Groups T-N and N-N at 0 sec. is not quite significant ($t = 1.859$, 18 *df*, $.10 > p > .05$). However, when the data from Groups T-T, T-N, and N-T at 0 sec. are pooled (*Ss* in these Groups are responding under essentially the same amount of tension in this immediate recall condition) and compared with Group N-N, the difference is significant ($t = 2.365$, 38 *df*, $p < .01$). The difference between Groups T-N and N-N at 240 sec. is not significant. The difference between pooled Groups T-T and N-T and pooled Groups T-N and N-N at 240 sec. is found to be significant at the .01 level of confidence ($t = 2.869$, 38 *df*).

Discussion

The obtained data show that as the time allowed for relaxation of tension induced during learning is increased, recall performance declines to the level obtained when no tension is induced during learning. This decline in performance is gradual, coinciding with the no-tension performance level after approximately 4 min. Davis has stated that 4 min. is the length of time needed for the dissipation of induced dynamogenic tension.[2] Thus, it appears that as tension dissipates, performance degrades. This result, coupled with the fact that performance does not decline when tension is induced during the recall trial, indicates that tension facilitates response elicitation. It is obvious from the present data that there is no difference between the T-T group and the N-T group. This fact also substantiates the conclusion that induced tension does not act upon the learning process. Induction during recall only is as facilitative as induction during both learning and recall. Thus, these results directly confirm Meyer's prediction. Tension might have produced an increase in magnitude or a decrease in latency of the responses but it clearly did not affect learning.

The data can be conceptualized in terms of S-R theory if the effect of tension is visualized as contributing to drive. For instance, in Hull's theory, ${}_sE_R$ is a multiplicative function of ${}_sH_R$ and drive. Since ${}_sH_R$ is not changed while ${}_sE_R$ is enhanced, tension must increase the drive.

Summary

Recall of paired adjectives was studied as a function of both temporal conditions of induced tension and time between learning and recall. Performance data were obtained for four groups, ten *Ss* in each, which differed according to whether tension was induced during learning, recall, both learning and recall, or neither learning nor recall. Time intervals between the learning and the recall trials were 0, 30, 60, 120, and 240 sec.

The results indicate that tension facilitates response elicitation but does not alter habit strength. Recall performance is no better if learning takes place under tension than if

[2] R. C., Davis, Personal communication, July 7, 1954.

learning takes place under no tension, provided enough time is allowed for relaxation. On the other hand, recall under tension is consistently better than recall under no tension. The results are in accord with the concept of tension as a drive.

References

1. Bills, A. G. The influence of muscular tension on the efficiency of mental work. *Amer. J. Psychol.*, 1927, **38**, 227–251.
2. Courts, F. A. Relations between experimentally induced muscular tension and memorization. *J. exp. Psychol.*, 1939, **25**, 235–256.
3. Courts, F. A. Relations between muscular tension and performance. *Psychol. Bull.*, 1942, **39**, 347–367.
4. Courts, F. A. The influence of practice on the dynamogenetic effect of muscular tension. *J. exp. Psychol.*, 1942, **40**, 504–511.
5. Davis, R. C. Patterns of muscular tension during "mental work" and their constancy. *J. exp. Psychol.*, 1939, **24**, 451–465.
6. Freeman, G. L. The optimal muscular tensions for various performances. *Amer. J. Psychol.*, 1938, **51**, 146–150.
7. Haagen, C. H. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, **27**, 453–463.
8. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
9. Meyer, D. R. On the interaction of simultaneous responses. *Psychol. Bull.*, 1953, **50**, 204–220.
10. Russell, J. T. Relative efficiency of relaxation and tension in performing an act of skill. *J. gen. Psychol.*, 1932, **6**, 330–343.

(Received August 23, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

STIMULUS GENERALIZATION FOLLOWING DIFFERENT METHODS OF TRAINING[1]

DANIEL B. REINHOLD[2] AND CHARLES C. PERKINS, JR.

Kent State University

Although stimulus generalization or some such principle seems to be accepted by most learning theorists, relatively few data are available which show how the shape and/or slope of the generalization gradient is influenced by experimental conditions. These limited data indicate that the slope of the gradient may be affected by the number of reinforcements (**7**), the proportion of training trials reinforced (**8**), strength of motivation (**1**), and type of motivation (**11, 12, 13**). The study reported here was designed to determine if the slope of the generalization gradient is steeper following a differential conditioning procedure in which the negative stimulus differs from the positive on a dimension other than the one along which generalization is tested, than following nondifferential training.

More specifically, *S*s in Group D were reinforced for running to one stimulus complex (e.g., black, smooth), but not for running to another stimulus (black, rough) differing along one dimension (roughness). Response strength was then tested to the positive training stimulus (black, smooth) and to a novel stimulus differing only along a second (color) dimension (white, smooth). There were also two control groups. For one, Group C, the procedure was the same as for Group D except that there were no nonreinforced trials to the negative training stimulus. The procedure for Group INT differed from that for Group D only in that all training trials, whether reinforced or not, were to the same "positive" stimulus.

METHOD

Apparatus.—An elevated runway 1.5 in. wide, 33 in. long, and 36 in. above the floor was used. The white starting box was 7.5 × 7 × 6 in. (inside measurement). Two separate sets of a runway, goal box, and goal-box stand were used. The reversible runways were constructed of hard wood covered on one side by .25-in. hardware cloth. The glass-covered goal boxes were 17 × 6.5 × 6.5 in. The two sets of these sections differed only in color—one was a near white (74% reflectance) and the other a near black (6% reflectance).

Lighting was provided by a 60-w. bulb 2.5 ft. above the center of the runway. The entire apparatus was surrounded by black and white cloth curtains which also served as one-way screens. The curtains were so arranged that when *S* was on the runway the background behind the goal box was half black and half white; the background to the starting box was entirely black. Thus, the two ends of the apparatus were presumably highly discriminable.

Timing was automatic; the clock started when the door of the starting box was opened and stopped when *S* crossed an electric-eye beam .25 in. in front of the goal box.

Subjects.—The *S*s were 12 male and 12 female experimentally naive albino rats approximately 90 days old at the beginning of the training.

Preliminary training.—All *S*s were first placed on a 24-hr. feeding schedule and reduced to 85% ad libitum weight. Throughout the study *S*s were kept at 85% weight by feedings immediately after the experimental period. The *S*s were run in squads of six, two in each group.

On the first day of preliminary training each *S* was placed on the runway (positive stimulus condition) and left there for 10 min. with the doors to the starting and goal boxes closed. Then *S* was moved to the positive goal box where it remained for 5 min. On Day 2 *S* received five trials on which it was placed in the middle of the runway and removed from the goal box after reinforcement, consumption of two .2-gm. pellets of Purina Laboratory Chow. On

[1] This paper is an abridgment of the first author's Master's thesis. The second author, who was the thesis director, is primarily responsible for preparation of this report.

[2] Now at Western Reserve University.

TABLE 1

SUMMARY OF THE EXPERIMENTAL PROCEDURE

Days	Group D (N = 8)	Group C (N = 8)	Group INT. (N = 8)
1–4	Preliminary training	Preliminary training	Preliminary training
5	4+ trials to positive stimulus 4− trials to negative stimulus	4+ trials to positive stimulus	4+ trials to positive stimulus 4− trials to positive stimulus
6–13	2+ trials to positive stimulus 6− trials to negative stimulus	2+ trials to positive stimulus	2+ trials to positive stimulus 6− trials to positive stimulus
14–16	4− trials to positive stimulus	4− trials to positive stimulus	4− trials to positive stimulus
(Test)	4− trials to novel stimulus	4− trials to novel stimulus	4− trials to novel stimulus

each of Days 3 and 4 all *Ss* had 10 reinforced trials from the starting box. The *Ss* were then assigned to the groups in a way which roughly equated them for running speed on Day 4.

Training.—As shown in Table 1, *Ss* in Group D received four reinforced trials to the positive stimulus (e.g., black, smooth) and four nonreinforced trials to the negative stimulus (e.g., black, rough) on Day 5. On Days 6–13 inclusive, these *Ss* were given two reinforced runs to the positive stimulus and six nonreinforced runs to the negative stimulus. The sequence of reinforced and nonreinforced trials was irregular.

On training days *Ss* in Group C received the same number of reinforced trials to the positive stimulus as *Ss* in Group D. However, they received no nonreinforced trials. The *Ss* in Group INT received the same sequences of reinforced and nonreinforced trials as *Ss* in Group D, but all trials were to the positive stimulus. Thus, Group INT received intermittent reinforcement.

Test.—On each of Days 14–16, all *Ss* were given eight test trials, four nonreinforced trials to the positive stimulus (P) and four nonreinforced trials to a novel stimulus (N) differing in color (e.g., smooth, white).

Counterbalancing.—The positive stimulus conditions differed for each of the four squads (of two *Ss* in each group) so that each combination of black or white, and smooth or rough was the positive stimulus for one squad. The negative stimulus (for Group D) always differed from the positive only in roughness, while the novel test stimulus always differed from the positive stimulus only in color. Possible extinction effects during test trials were counterbalanced by using a PNNPNPPN sequence for half the *Ss* and a NPPNPNNP for the other half.

RESULTS

The response measure used throughout is the common logarithm of *S*'s response time, the interval between the opening of the starting-box door and the breaking of the electric-eye beam as *S* started into the goal box.

On some trials *S* did not respond within the 30 sec. permitted; therefore, means could not be used. All group measures were obtained by getting the median log response time for each *S* on the designated trials and then obtaining the group median of these medians. All levels of confidence are based on the *U* test (**10**). Probability values are expressed in terms of a two-tailed test even though most differences were predicted and one-tail p values are given in Mann and Whitney's tables.

On the last day of training all three groups had about the same median median log response time to the positive stimulus. The values were .23, .27, and .33 sec. for Groups D, INT, and C, respectively. No difference between groups on this day was significant at the 10% level of confidence.

Apparently, a discrimination had been established in the *Ss* of Group D since they all had a shorter median latency to the positive than to the negative stimulus on the last day of training.

Generalization gradients based on measures obtained on the first test day are presented in Fig. 1. Since every *S* had a shorter median log response

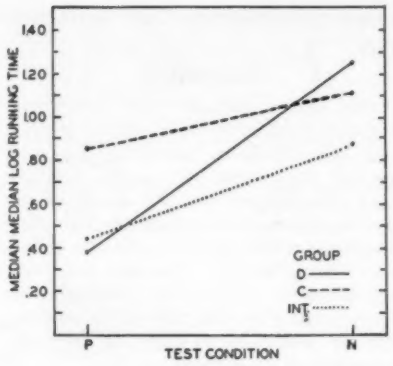

FIG. 1. Generalization gradients for each group based on running time on the first day. Any given point was obtained by first deriving the log of the median running time for each *S* and then determining the median of these logs for the eight *S*s tested to the positive training stimulus (P) and a novel stimulus (N). Greater response strength is indicated by the lower point.

time to the positive training stimulus (P) than the novel stimulus introduced on test trials (N), the generalization decrements can not be attributed to sampling errors. The gradient is steeper, the decrement greater for Group D than for either of the other groups.

In order to test the significance of the difference between the size of the decrement for Group D and the other two groups, the median log response time to Stimulus P on the first test day was divided by the median log response time to Stimulus N. A U of 1 ($p < .001$) was obtained in testing the difference between these ratios for Groups D and C. The U for the ratios of Groups D and INT was 14 ($p < .07$).

A second index of the size of the generalization decrement for individual *S*s was median log response time to Stimulus N minus median log response time to Stimulus P. For this index the U between Groups D and C is 2 ($p < .001$). A similar comparison of Groups D and INT yields a U of 11 ($p < .03$).

The greater decrement for Group D than Group C is highly significant with both indices. Since the gradients for these groups cross and since the two indices (the first giving a decrement relative to response strength to the positive stimulus, and the second index measuring the absolute decrement) give the same results, it may safely be concluded that the generalization gradient is steeper for Group D than for Group C. The difference between the decrements for Groups D and INT is not as reliable. However, since both indices give borderline significance and the gradients cross, it seems highly probable that the difference is not merely the result of chance or the selection of inappropriate units of measurement.

For the first index the U for Groups C and INT is 12 ($p < .04$); for the second index U equals 16 ($p < .11$). The difference in size of decrement for the two control groups, while highly suggestive, is of borderline significance. Also, since the two gradients do not cross, the difference in the size of the decrements is dependent upon the units of measurement selected. Nevertheless, it is of interest that the difference in decrement size for the two control groups is in the opposite direction from the results obtained by Humphreys (**8**), who found the gradient to be steeper following 100% reinforcement than intermittent reinforcement.

When the data are plotted for all three test days, the gradients are quite similar to those obtained on the first test day except that differences between the groups are somewhat less marked and that the gradients for the two control groups are almost exactly parallel.

The median median log running

time to each test condition for each group on the three test days, is presented in Table 2. Examination of this table disclosed that for Group C there appears to be extinction of the response to both Stimulus P and Stimulus N. However, for Groups D and INT there is marked extinction of the response to Stimulus P but no tendency towards its extinction to Stimulus N. Since every *S* in both groups (D and INT) had a longer median log running time to Stimulus P on test Day 3 than on test Day 1 while just half the *S*s in each of these two groups showed an increase in median log running time to Stimulus N, it seems unlikely that this difference in extinction rates is merely a chance phenomenon. The more rapid extinction of the response to the training stimulus than to the novel test stimulus in Groups D and INT indicates a flattening of the generalization gradient during test trials. Thus, for Group D the generalization decrement is .87 on Day 1 and only .10 on Day 3; for Group INT the decrement drops from .43 to .06 from Day 1 to Day 3. This finding is in line with results of Grant and Schiller (**4**), but is the opposite of the results obtained by several other investigators (**1, 7, 9**). Further work is clearly needed to determine more about the apparently complicated relationship between generalization and extinction rate.

TABLE 2

MEDIAN MEDIAN LOG LATENCY FOR EACH GROUP UNDER EACH TEST CONDITION ON TEST DAYS 1, 2, AND 3

Group	Positive Stimulus			Novel Stimulus		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
D	.38	.60	1.04	1.25	1.24	1.14
C	.86	1.21	1.17	1.11	1.32	1.41
INT	.44	.61	.77	.87	.99	.83

DISCUSSION

Our results indicate that differential training along one dimension may steepen the generalization gradient along a second dimension. It seems unlikely that this finding was the result of the development of orienting or attending responses which would make the difference between the training and test stimuli more distinctive. Attending to the primarily tactual cues of the runway surface would be unlikely to facilitate reception of color cues.

The demonstrated steepening of a generalization gradient following differential training along another dimension may be loosely thought of as the development of a set to discriminate and thus seems to be closely related to the development of learning sets (**5**).

The steepening effect of such differential training presumably affects the generalization gradients obtained following a classical conditioning procedure where differential conditioning is necessarily involved (**14, 15**). In fact, inter-experimental comparison of generalization following classical conditioning, where clear decrements were obtained (e.g. **6, 9**), with the horizontal gradient obtained by Ferster (**2**), following non-differential training, led to the design of our study. Such comparisons suggest that our major finding is a general phenomenon not confined to the specific conditions employed in this experiment. Thus, our results emphasize the need for a clear distinction between differential and nondifferential training by indicating another way in which they have different consequences.

Further work is needed to determine how generalization along one dimension is affected by the amount of differential training along another dimension, the similarity of positive and negative training stimuli, and related factors.

SUMMARY

Three groups of eight rats each were trained to traverse a runway. The training procedure

involved differential reinforcement of the running response to presence or absence of hardware cloth (Group D), 100% reinforcement of the response to the positive stimulus (C), or intermittent reinforcement (INT). All groups received identical nonreinforced test trials to the positive training stimulus and to a novel stimulus differing in color.

A significant generalization decrement was obtained for all groups; i.e., response strength was less to the novel stimulus than to the positive training stimulus. The generalization decrement was significantly ($p < .001$) greater for Group D than Group C. The decrement for Group D was also greater than for Group INT, but this difference was only of borderline significance.

It was concluded that differential conditioning along one dimension increases the slope of the generalization gradient along a second dimension.

References

1. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.
2. Ferster, C. B. The effect on extinction responding of stimuli continuously present during conditioning. *J. exp. Psychol.*, 1951, **42**, 443–449.
3. Graham, C. H., & Gagne, R. N. The acquisition, extinction, and spontaneous recovery of a conditioned response. *J. exp. Psychol.*, 1940, **26**, 251–280.
4. Grant, D. A., & Schiller, J. J. Generalization of the conditioned galvanic skin responses to visual stimuli. *J. exp. Psychol.*, 1953, **46**, 309–313.
5. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
6. Hovland, C. I. The generalization of conditioned response: I. The sensory generalization of conditioned responses with varying frequency of tone. *J. gen. Psychol.*, 1937, **17**, 125–148.
7. Hovland, C. I. The generalization of conditioned response. IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 161–176.
8. Humphreys, L. G. Generalization as a function of the method of reinforcement. *J. exp. Psychol.*, 1939, **25**, 361–372.
9. Littman, R. A. Conditioned generalization of the galvanic skin reaction to tones. *J. exp. Psychol.*, 1949, **39**, 868–882.
10. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
11. Miller, N. E., & Kraeling, D. Displacement: Greater generalization of approach than avoidance in a generalized approach-avoidance conflict. *J. exp. Psychol.*, 1952, **43**, 217–221.
12. Miller, N. E., & Murray, E. J. Displacement and conflict: Learnable drive as a basis for the steeper gradient of avoidance than of approach. *J. exp. Psychol.*, 1952, **43**, 227–231.
13. Murray, E. J., & Miller, N. E. Displacement: Steeper gradient of generalization of avoidance than of approach with age of habit controlled. *J. exp. Psychol.*, 1952, **43**, 222–226.
14. Perkins, C. C., Jr. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46**, 225–231.
15. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.

(Received August 24, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

STIMULUS AREA, STIMULUS DISPERSION, FLASH DURATION, AND THE SCOTOPIC THRESHOLD[1]

OSCAR S. ADAMS, DAVIS J. CHAMBLISS,[2] AND ARTHUR J. RIOPELLE

Emory University

It is well known that the intensity of stimulation required to evoke a scotopic threshold response is dependent, among other things, on the area and duration of the stimulating flash. Numerous studies (**2, 3, 4, 8, 10**) have shown that with an increase in area there is, within limits, a subsequent decrease in the intensity necessary to evoke a threshold response. The exact function of this reciprocity is still undetermined. Within the critical duration, it is known (**4, 7, 8**) that a reciprocity exists between time and intensity (Bunsen-Roscoe Law). When the area, time, and intensity variables are combined the result is that there is an interaction between area and the relation described by the Bunsen-Roscoe Law. The transition from $It = k$ to $I = k$ is more abrupt with small areas than with large ones (**4, 8**).

In addition to the time-area-intensity relations, the absolute scotopic threshold is a function of the separation between two or more noncoincident but equally illuminated retinal areas. Beitel (**1**), extending previous work on spatial summation of the CFF (**5, 6**), has found that increasing the separation between two constant area patches produces first a rapid, then a slow, and later a rapid increase in the threshold. Riopelle (**9**) has studied the effect of dispersion of area into several smaller subareas, and has found an increase in threshold as a function both of the degree of dispersion as well as the size (or number) of the stimulus subareas. In each condition studied the total area remained constant.

The present investigation was designed to study spatial summation as a function of the time-area-intensity relations. Specifically, it is parametric in nature, intended to (*a*) survey the extent of summation which occurs as the separation is increased between two patches equal in size and illumination, (*b*) to investigate the effects of variations of exposure time and area on this extent, and (*c*) to compare the form of the resultant family of curves.

METHOD

Subjects.—The three writers served alternately as *S* and *E*. All were well acquainted with the task at hand.

Apparatus.—The apparatus has been previously described (**10**), and a schematic diagram is shown in Fig. 1. The essential feature of it was rigid control of the intensity and duration of the test flash. The *S* was seated in a small light-tight cubicle built into one corner of the experimental room. The opal glass screen was permanently mounted against an opening in the cubicle wall which faced *S*.

Stimuli.—Three series of black-painted metal masks were used between the opal glass screen and the opening in the cubicle. Series I consisted of 14 masks, of which 12 had two ¼-in. holes separated vertically from a common point. Mask 13 had a single ¼-in. hole whose center coincided with the common point of separation, and Mask 14 also contained a single hole,

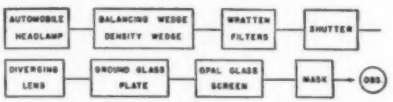

FIG. 1. Schematic diagram of the apparatus.

[1] This research was supported in part by a grant from the University Center in Georgia, Inc.

[2] Now at the University of Wisconsin.

similarly placed, with area equal to that of two ¼-in. holes. Series II consisted of 12 masks with a base diameter of ¼ in. The single and double area patches were used in this series and the following group, Series III, which consisted of 11 masks in which the stimulus patches were ⅛ in. in diameter. The separation ranged from adjacent to 3¼ in.; or in visual angle, the smallest angle subtended between centers was 14′, and the largest was 6° 26′.

Procedure.—Prior to determining threshold measurements, *S* was dark adapted for 40 to 45 min. A short period at the end of the dark adaptation period was spent presenting *S* with flashes in order for him to establish a stable threshold criterion. All judgments were made with the left eye which fixated a small point source of light located 8 in. to the left of the center of the opal glass screen. A forehead-and-chin rest was used to maintain a constant distance of 30½ in. from *S*'s eye to the stimulus mask. The test patch was thus imaged about 14° from the fovea on the temporal retina.

Absolute reappearance thresholds were taken in the following manner. For each threshold determination, *S* was alerted by a verbal signal, and then given two dim, but definitely supraliminal flashes in quick succession. The third flash was presented well below threshold intensity. The wedge was then moved about .25 cm. (about 1/80 log unit) toward the brighter end with each successive flash until *S* responded affirmatively. Two successive affirmative responses constituted a trial.

All flashes were presented at either 15 or 150 msec. duration, being, respectively, below and above the critical duration. The average of 20 determinations for each mask of the three series at each of the two durations was taken as *S*'s threshold for that mask at the given duration. Measurements were recorded in terms of log micro-microlamberts.

Results and Discussion

The results of the experiment are shown in Fig. 2 where the threshold in log micro-microlamberts is plotted against the distance between the centers of the stimulus patches. The parameters of the curves are the three stimulus areas and the two durations of exposure. The "D" and "S" points on the abscissa represent, respectively, the double and single area patches. There is a certain

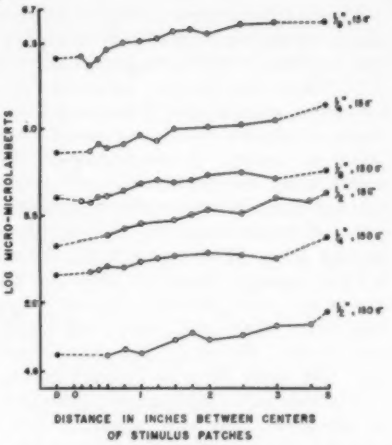

Fig. 2. Scotopic thresholds for the three series of test patches presented at 15 and 150 msec. duration.

logical difficulty involved in considering these points as belonging to the curves since they are single points; the curves are broken to indicate this. They are included primarily as check points, since it was not expected that even complete summation between two patches, within the range of experimental error, could yield a threshold more sensitive than a single patch equal in area to the two separated patches. Similarly, when the limits of summation are exceeded, a single patch should give about the same threshold as the two separated patches.

The data clearly show that for both flash durations and for all stimulus areas used in this experiment, an increase in the dispersion of the stimulus results in an increased threshold. Taken as a group, the curves are very near parallel. We conclude that when plotted in log micro-microlamberts, the effects of separation are independent of the effects of the other variables of area and duration.

Beitel (1) has reported that summation clearly existed up to about 150′ of separation, and that it was not apparent beyond this. Here the greatest separation for any series was at least 5° between inner edges, or slightly over 6° measured from center to center. A comparison of the points along the curves with the single patches indicates that there probably was not much summation taking place beyond 3° to 4° of separation. Beitel further reported that the form of his curves was sigmoidal, that is, that there was first a sharp rise, then a relatively flat portion, followed by a final sharp rise. No such form is evident in the curves obtained in this study. The lack of agreement on this point is very probably due to the difference in shape of the test patches. Where Beitel's patches were semicircular apertures, ours were circular. It seems obvious to expect greater summation effects at small separations with the former type patch, where the area of separation is bounded by more area of stimulation, than with the circular patch.

Summary

This study presents further evidence to support the well-established finding that an increase in area and exposure time causes a decrease in the absolute scotopic threshold. In addition, the data indicate that spatial summation is essentially independent of the variables of area and duration within the range used in the experiment.

References

1. Beitel, R. J., Jr. Spatial summation of subliminal stimuli in the retina of the human eye. *J. gen. Psychol.*, 1934, **10**, 311–327.
2. Graham, C. H. The relation between area and intensity of visual threshold. *Amer. J. Psychol.*, 1930, **42**, 420–422.
3. Graham, C. H., Brown, R. H., & Mote, F. A., Jr. The relation of size of stimulus and intensity in the human eye: I. Intensity thresholds for white light. *J. exp. Psychol.*, 1939, **24**, 555–573.
4. Graham, C. H., & Margaria, R. Area and intensity-time relation in the peripheral retina. *Amer. J. Physiol.*, 1935, **113**, 299–305.
5. Granit, R. Comparative studies on the peripheral and central retina: I. On interaction between distant areas in the human eye. *Amer. J. Physiol.*, 1930, **94**, 41–50.
6. Granit, R., & Harper, P. Comparative studies on the peripheral and central retina: II. Synaptic reactions in the eye. *Amer. J. Physiol.*, 1930, **95**, 211–228.
7. Hartline, H. K. Intensity and duration in the excitation of single photoreceptor units. *J. cell. comp. Physiol.*, 1934, **5**, 229–247.
8. Karn, H. W. Area and intensity-time relation in the fovea. *J. gen. Psychol.*, 1936, **14**, 360–369.
9. Riopelle, A. J. Dispersion of stimulus-areas and the scotopic threshold. *Amer. J. Psychol.*, 1951, **64**, 380–390.
10. Riopelle, A. J., & Chow, K. L. Scotopic area-intensity relations at various retinal locations. *J. exp. Psychol.*, 1953, **46**, 314–318.

(Received August 23, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

THE INFLUENCE OF OVERTONE STRUCTURE ON THE PITCH OF COMPLEX TONES[1]

WILLIAM H. LICHTE AND R. FLANAGAN GRAY[2]

University of Missouri

That pitch is a subjective phenomenon influenced, in pure tones, by all the physical characteristics of the stimulus is now widely understood. Little is known about the pitch of complex tones and the common assumption probably is that their pitch is almost entirely determined by the frequency of the fundamental. Effects of overtone structure (or spectrum) on pitch are indicated, however, in a few statements in the psychological literature. Banister (**1,** p. 894) says, referring to a complex tone, ". . . its pitch is usually slightly higher than that of a pure tone of the same frequency. The more pronounced the overtones, especially the higher ones, the higher the pitch appears to be." Banister gives no citation in connection with this statement, which may be based on his own subjective impressions. Helmholtz (**3,** p. 61) apparently made a similar judgment, for he says, ". . . although the pitch of a compound tone is, for musical purposes, determined by that of its prime, the influence of the upper partial tones is by no means unfelt. They give the compound tone a brighter and higher effect." These unsupported statements have apparently not made much of an impression. In the current handbook of experimental psychology, for example, Licklider (**6**) discusses the pitch of complex tones only in the case of the missing fundamental.

Variations in the pitch of complex tones, by subject and by tone, were found by Lichte (**5**) in a control study of a previous experiment. All *Ss* judged the pitch of two tones which were equal in fundamental frequency, but differed in spectrum. Three groups of eight *Ss* each judged pairs of tones at frequencies of 180 cps. In each group there were *Ss* for whom the two tones differed significantly in pitch, but the differences were not systematic. Significant individual differences in the pitch of a given complex tone were also found. A pair of tones at 360 cps and one at 540 cps were also judged by three *Ss* each. In every case the tone in which the lower partials were strong was *higher* in pitch than the tone in which the high partials were strong, and five of the six differences were statistically significant. These results contradict Banister's statement, but with six *Ss* it can only be said that a difference in this direction was characteristic of some individuals.

The statements and the data referred to suggest that the spectrum of tones does influence their pitch. The purpose of the present study was to verify the existence of this effect. The experimental hypothesis was that two complex tones with the same fundamental frequency but markedly different spectra will differ in pitch.

[1] This article was written while the first author held an appointment as Summer Research Fellow; the appointment was granted through the University Research Council, University of Missouri.

[2] Now at Aviation Medical Acceleration Laboratory, Naval Air Development Center, Johnsville, Pennsylvania.

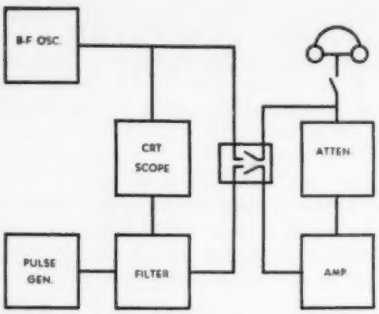

FIG. 1. Block diagram of the apparatus.

METHOD

Apparatus.—A block diagram of the apparatus is shown in Fig. 1. The basic source of the complex tones was a pulse generator (Measurements Corporation Pulse Generator, Model 79-B). As Licklider (**6**, p. 987) indicates, when the waveform of the stimulus is a series of pulses or spikes of very brief duration, a harmonic analysis shows that the stimulus actually consists of an infinite series of harmonics of equal intensity. This "raw material" was put through appropriate filters to produce sounds with different spectra. The output of the pulse generator and filter was led through an amplifier and an attenuator to headphones (Permoflux High Fidelity Earphones, Model PDR-8). The pure tone source was a beat-frequency oscillator (General Radio, Type No. 1304-A), which was connected to the same earphones. Both signals were taken through a "joy stick" toggle switch (dpdt) which *S* could move from side to side, connecting the earphones alternately to the two sound sources; when the switch was untouched, both connections were broken. Another switch enabled *E* to disconnect the earphones at will.

The pure tone and complex tone signals were also led to the horizontal and vertical plates, respectively, of a cathode-ray oscilloscope. The purpose of this was to enable *E* to set the pure and complex tones at the same frequency. The two signals produced a Lissajous figure on the scope face; this rotated when the frequencies were unequal and could be made motionless by adjusting the frequency of either tone until they were equal.

Stimuli.—At each of two frequencies, 250 cps and 700 cps, two complex tones were used. In one tone of the pair the high partials were present and were relatively strong; in the other tone only the lower partials were present. The spectra of the four tones are shown in Fig. 2. The tone with a fundamental frequency of 250 and relatively strong low partials was designated 250-L, the tone with the same frequency and

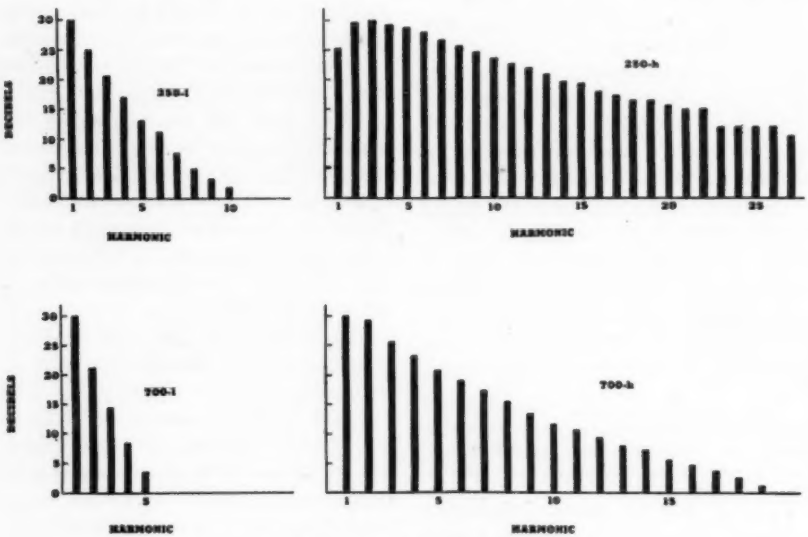

FIG. 2. Spectra of the complex tones.

strong high partials was called 250-H, and those at the higher frequency were called 700-L and 700-H. The spectra of the tones were determined from measurements made with a wave analyzer (General Radio, Type No. 736-A).

Subjects.—The *S*s were volunteers from the beginning psychology courses at the University of Missouri. After volunteering, they were used only if they scored above the 30th percentile on the Pitch Test (Form A) of the Seashore Measures of Musical Talents. Of those accepted, six were later rejected because they could not make judgments within the range of the incremental frequency dial of the oscillator (± 50 cycles). The *S*s were randomly assigned to the two groups as they were accepted. There were 26 *S*s in Group I (which judged tones 250-H and 250-L) and 23 *S*s in Group II (which judged tones 700-H and 700-L).

Procedure.—The apparatus was allowed to warm up for 30 min. before the beginning of each session. At the beginning of the session the oscillator was taken as a frequency standard and the pulse generator was set at the desired frequency (250 or 700 cps) by means of the Lissajous figure. At the start of each series of judgments on a tone *S* adjusted the intensity of the complex tone until it was equal in loudness to the pure tone (which was held constant at an intensity level of approximately 65 db).

The task of *S* on each trial was to adjust the incremental frequency dial of the oscillator until the pure tone was equal in pitch to the complex tone. The setting of the incremental frequency dial plus the standard frequency was then taken as a measure of the pitch of the complex tone. One *S* might judge the 250 cps complex tone to be equal in pitch when the dial was at + 7 (oscillator at 257 cps). For convenience, the pitch of the complex tone would then be recorded as + 7. After each five trials the frequency calibration was checked and was re-done if necessary (this was rarely needed).

Each *S* served in three periods of about 40 min. each. The first was a practice session, the latter two were the regular experimental sessions (1 and 2). In the practice session *S* was first given the Seashore test. After this, the first practice task was to adjust the intensity of one of the complex tones until it was equal to the pure tone in loudness. Then the nature of the pitch judgment was explained and *S* made ten settings of equality. He then set the other complex tone equal in loudness to the pure tone and made ten pitch judgments for it. In making a pitch equation, *S* was allowed to alternate the tones as many times and at whatever rate he chose.

In each of the following sessions (1 and 2), *S* made a loudness equation and 25 pitch judgments on each of the two complex tones in the pair. The order of tones was counterbalanced and in each group about half the *S*s used each order. The *E*'s pre-settings of the incremental frequency dial were taken from a table of random numbers. An equal number of plus and minus settings was used and numbers from 0 to 30 in the table were used.

The *S* sat with his arms resting on a table and manipulated the toggle switch with his left hand and the incremental frequency dial with his right. A cloth screen prevented his seeing the incremental frequency dial. The face of the oscilloscope was not visible to *S*, and on all calibrations *E* opened his switch so that *S* could never accidentally hear the tones.

TABLE 1

GROUP MEANS AND *SD*'S (IN CYCLES PER SECOND) OF EQUAL-PITCH SETTINGS FOR GROUPS I AND II

Group	*N*	Tone	Mean	*SD*
I	26	250-L	5.88	13.85
I	26	250-H	13.43	17.47
II	23	700-L	6.39	19.74
II	23	700-H	16.47	29.73

RESULTS

All data for the pitch matchings are given in cycles per second above or below equality with the fundamental frequency of the complex tone.[3] The group means and *SD*'s for each tone are given in Table 1. Each mean gives the frequency of the pure tone when the frequency of the complex tone is taken as zero. On the average, each complex tone was heard as equal in pitch to a pure tone with a frequency higher than its fundamental. The size of the *SD*'s indicates important individual differences; these are apparent in the four curves of Fig. 3, which gives the distributions of individual means on each tone. To take an extreme example, one *S* heard tone 700-H as equal in pitch to

[3] Detailed data on the results have been deposited with the American Documentation Institute. Order Document No. 4519, remitting $1.25 for microfilm or $1.25 for photocopies.

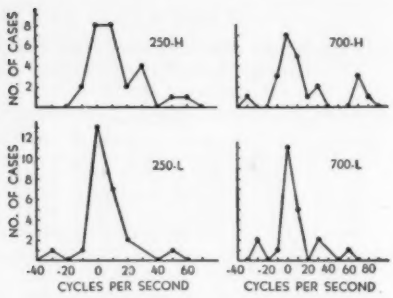

FIG. 3. Distributions of individual means for each tone.

a pure tone 38.8 cps lower in frequency while another *S* heard it as equal to a pure tone 79.9 cps higher in frequency. These graphs show the modes to be at the class interval with a mid-point of zero; most of the remaining means are in the plus range, and a few of these are rather high.

There was also a wide variation in the *SD*'s of the 50 judgments of each *S* on each tone. The highest was 43.7 and, on the other hand, one *S* obtained an *SD* of .14 on each of the two complex tones. The large *SD* came from an atypical *S* who gave all minus settings on that tone in the first session and all plus settings in the second session. The small *SD* was the result of 49 settings of *0* and one of *+ 1* on each tone. For the combined groups, 40% of the *SD*'s were below 5.0 and 75% were below 15.0. At 700 cps these *SD*'s were frequently higher for the H tone than for the L tone. This finding and the *SD*'s shown in Table 1 suggest that the pitch of the H tones was harder to judge than that of the L tones.

The H tone was judged higher in pitch than the L tone by each group. The differences between tones, *SE*'s, and *t*'s are shown in Table 2. The *SE*'s were computed directly from the differences, *S* by *S*, between the means for the two tones. Both of the group

TABLE 2

GROUP DIFFERENCES BETWEEN TONES, *SE*'S, AND *t*'S FOR EACH GROUP

Group	*N*	Diff.$_{(H-L)}$	SE_d	*t*
I	26	7.55	1.96	3.85
II	23	10.08	2.88	3.50

differences were significant at less than the 1% level of confidence.

The distributions of differences are shown in Fig. 4. At 250 cps all the differences were in one direction (the H tone higher) or were zero; at 700 cps the H tone was *lower* in pitch for a few *S*s. For a few *S*s the tones differed in pitch by 35–40 cps. In Group I the tones differed significantly in pitch for 16 of the 26 *S*s, in Group II for 15 of 23 *S*s (5% level of confidence).

Because each *S* served in two sessions the session-to-session differences in group means were obtained. Two of these were plus and two were minus and the largest (3.08 cycles) was not significant at the 10% level of confidence.

DISCUSSION

Research workers in the sensory fields often assume that the range of individual

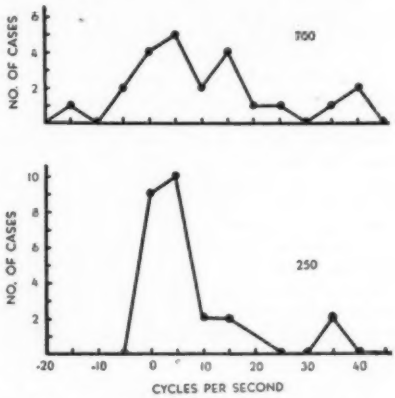

FIG. 4. Distributions, by individual, of differences between paired tones.

differences in basic sensory processes is rather narrow. To the writers the range of individual differences on each of the four tones seems to be an important indication of the complexity of the pitch process. Differences in past experience were perhaps the basis for the individual differences in the pitch of these tones.

The basic finding to be evaluated is the difference between the H and L tones. It should first be pointed out that the differences in spectrum between the H and L tones were not highly systematic and the difference between the 250 cps tones was not the exact equivalent of the difference between the 700 cps tones.

Two points will help in evaluating the size of the pitch differences between the H and L tones. The obtained differences are three to four times the size of the frequency DL's for pure tones at these frequencies. In terms of the musical scale, the difference at 250 cps is equal to .26 whole tones; that at 700 cps is equal to .12 whole tones. These differences are large enough to be important from the standpoint of musical performance.

Probably the most important problem in the evaluation of these results is the possible influence of intensity on the pitch of complex tones. For all tones and all *S*s, it should be remembered, the pure tone was at a constant intensity level. Each *S*, however, set the intensity of the complex tone to make it equal in loudness to the pure tone. Differences in these settings might have produced the pitch results if intensity does influence pitch.

A study by Lewis and Cowan (**4**) seems to indicate that intensity does not influence the pitch of complex tones on the violin and 'cello, but the validity of the study depends upon the assumption that the direction and magnitude of such effects would vary with frequency in the same fashion as those for pure tones (with frequencies equal to the fundamentals of the complex tones). This does not seem to the writers to be a satisfactory assumption. An instance in which intensity did affect pitch is described by Fletcher (**2**). A 200-cps tone with five components changed in pitch by .17 whole tones when the loudness level was raised from 40 db to 100 db. In the present study, the variation in the intensity of the complex tones resulting from *S*s' equal-loudness settings was certainly much smaller than 60 db, and would therefore yield pitch variations (assuming Fletcher's case is typical) much smaller than .17 whole tones. The obtained finding of .26 whole tones at 250 cps is thus well beyond the possible effect of any intensity changes.

If the intensity of complex tones did influence their pitch, a relationship should be found in the data of the present study between intensity level and mean pitch for the complex tones. The rank-order correlations between these two variables were obtained and are shown in Table 3. (Rank order correlations were necessary because the intensity dial on the amplifier showed arbitrary units of unknown value.) Most of these correlations were low and only one of them (for tone 700-H on Session 1) was significant at the 5% level of confidence. These cannot be interpreted as giving a final answer, but it does seem probable that the results of the experiment were not due to the influence of loudness on pitch.

It is not easy to reconcile the data of the present study with those of the previous study by Lichte (**5**). In the earlier study *no* systematic differences were found at 180 cps, whereas at 360 and 540 cps the obtained differences were in the opposite direction (tones with strong high partials were lower in pitch) from those of the present study. The differences in the spectra of the stimuli for the two studies are im-

TABLE 3

RANK ORDER CORRELATIONS BETWEEN INTENSITY AND PITCH OF COMPLEX TONES

Tone	*N*	Session 1	Session 2
250-L	26	.02	.25
250-H	26	.35	.04
700-L	23	−.37	−.15
700-H	23	−.49*	−.03

* $p = .05$.

portant. None of the stimulus pairs at 180 cps in the early study was closely similar in spectrum to those of the present study. The stimulus pairs at 360 and 540 cps contained only eight and five partials, respectively, whereas in the present study tone 250-H contained over 25 partials and tone 700-H contained 19. In view of these differences in stimuli, both as to spectrum and frequency, similar results are hardly required. Furthermore, the results of the early study are less reliable because of the smaller number of *Ss*. The two studies do agree in showing individual differences in pitch for any given complex tone.

The present results support Banister's statement (previously quoted) completely and are probably in agreement also with the effect Helmholtz described.

Although the influence of the intensity of complex tones on their pitch deserves investigation, the writers feel that the present study does show that the pitch of complex tones is influenced by spectrum. This fact seems to be of theoretical and practical importance. To explain the moderately large change in the pitch of a 250 cps tone when the periodicity of the wave does not change and when the point of greatest magnitude of response on the basilar membrane probably does not change is an added burden for either the frequency or the place theorist. From the practical standpoint, the findings are significant with regard to musical performance. Precision of intonation is much finer than .26 (or .12) whole tones; thus the musician must, whether he knows it or not, take account of this effect in performance.

Summary

The pitches of pairs of complex tones at fundamental frequencies of 250 and 700 cps were measured; one tone of each pair contained only the lower partials, the other contained many strong high partials. Pitch was measured by having the *Ss* adjust the frequency of a pure tone until it equalled the complex tone in pitch. The pure and complex tones were equally loud. The *Ss* were university students who ranked in the upper 70% on the Seashore pitch test. Twenty-six *Ss* judged the pitch of the tones at 250 cps, 23 *Ss* judged the tones at 700 cps.

Marked individual differences in pitch were found for each complex tone. At each frequency, the tone with the strong high partials was higher in pitch. The difference was equivalent to 8 cycles (.26 whole tones) at 250 cps and 10 cycles (.12 whole tones) at 700 cps.

The results did not seem to be caused by variation in the intensity settings of the complex tones. The overtone structure of a complex tone appears to influence its pitch to a degree that is of theoretical and practical significance.

References

1. Banister, H. Auditory phenomena and their stimulus correlates. In C. Murchison (Ed.), *A handbook of general experimental psychology*. Worcester, Mass.: Clark Univ. Press, 1934.
2. Fletcher, H. Loudness, pitch and the timbre of musical tones. *J. acoust. Soc. Amer.*, 1934, **6**, 59–69.
3. Helmholtz, H. (Trans. by A. J. Ellis.) *Sensations of tone*. 3rd Trans. Ed. of 4th German Ed. London: Longmans, Green, 1895.
4. Lewis, D., & Cowan, M. The influence of intensity on the pitch of violin and 'cello tones. *J. acoust. Soc. Amer.*, 1936, **8**, 20–22.
5. Lichte, W. H. Attributes of complex tones. *J. exp. Psychol.*, 1941, **28**, 455–480.
6. Licklider, J. C. R. Basic correlates of the auditory stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.

(Received August 5, 1954)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

ACCURACY OF POSITIONING RESPONSES AS A FUNCTION OF SPRING LOADING IN A CONTROL[1]

HARRY P. BAHRICK
Ohio Wesleyan University
WILLIAM F. BENNETT AND PAUL M. FITTS
The Ohio State University

A rationale for predicting the manner in which control loadings can augment proprioceptive information arising from the movement of a control has been developed in another paper (**2**). The predictions regarding different types of control loadings are that (*a*) increasing the spring constant of a control will increase the accuracy with which its position is discriminated, (*b*) increasing its viscous damping will increase the accuracy with which its rate of motion is discriminated, and (*c*) increasing its mass will increase the accuracy with which its acceleration is discriminated. The present paper attempts to test the prediction regarding the effects of spring loading upon the accuracy of positioning responses.

When a control is spring loaded, the force required to maintain it in any given position is directly proportional to the displacement of the control with respect to some reference position. It is assumed that force can be used as a cue for a positioning response, and that the usefulness of this cue will vary as a function of the degree of discriminability among the forces associated with different positions. It is therefore hypothesized that the degree to which different spring loadings of a control effect accuracy in positioning a control is a function of the discriminability of the forces determined by a particular type of loading.

Data regarding the accuracy of applying pressures to a semi-rigid control stick are available from a study by Jenkins (**3**) who used a method of average error. He found that the relative *dl* decreased appreciably from 1 to 10 lbs. and was approximately constant at about .06 beyond 10 lbs. From these data one can make two predictions regarding the comparative accuracy expected in the positioning of controls using various types of force displacement relations. (*a*) Positioning errors should become smaller as the ratio $\frac{\Delta I}{I(\Delta D)}$ is increased, where I is the force required to achieve the specified displacement, and ΔI is the force change associated with a given displacement change, ΔD. For example, greater positioning accuracy should be achieved with a control in which an error of 1 in. in positioning the control is associated with a force change from 9 to 10 lbs. than with a control in which a 1-in. error is associated with a force change from 19 to 20 lbs. (*b*) For any given ratio of ΔI to $I(\Delta D)$, positioning errors should be less as the value of I is increased from 0 to 10 lbs., and errors should be relatively con-

[1] This research was conducted in the Laboratory of Aviation Psychology of The Ohio State University through support by the United States Air Force under Contract No. AF 33(038)-10528 with The Ohio State University Research Foundation, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, publication, use, and disposal in whole or in part by or for the United States Government.

stant for values of I above 10 lbs. For example, a control movement in which a positioning error of 1 in. is associated with a force change from 8 to 10 lbs. should be made with greater accuracy than a movement in which the same displacement error is associated with a force change from 4 to 5 lbs.

These predictions must be qualified by several other considerations. Jenkins' data were obtained with a control which was essentially isometric and thus the arm remained essentially fixed during the response. Effective pressure stimulation while the arm is in motion, however, is determined not only by the elasticity constant of the control, but by such other factors as (*a*) the mass and damping of the control, (*b*) the physical constants of the limb, and (*c*) the changing effect of these variables on receptors in the arm as a result of the changing geometry of forces during movement of the limb.

Orlansky's (**4**) conclusion that spring-loaded controls should be used in aircraft is in conformance with the present predictions. However, empirical tests of the effect of various spring constants upon accuracy of positioning movements have failed to confirm the above predictions in several instances. Weiss (**5**) found no significant effect of the spring constant of a control upon constant or variable errors of positioning responses, and concluded that distance is a more important cue than pressure in positioning motions. However, in Weiss' experiment the ratio of control displacement to *relative* pressure change was not evaluated, and visual information regarding the extent of control motion was not excluded. In another study (**2**) changes in spring stiffness failed to produce significant changes in the accuracy of spatial reproduction of simple movements. However, in this study there was little opportunity for learning since no knowledge of results was provided and only seven trials were given to each *S*.

In the present study the relation between displacement and relative as well as absolute changes in torque of a spring-centered control was varied and the effect upon accuracy of positioning movements determined. The design permits an evaluation of the relative importance of cues associated with amplitude of movement per se, of terminal torque per se, and of the rate of change of torque with amplitude. Knowledge of results was provided during extended practice and all visual cues were eliminated.

Method

Apparatus

A schematic diagram of the apparatus is shown in Fig. 1. Two types of spring loading were used. With one type tension was made to increase rapidly in any desired magnitude as the control was moved, through the use of interchangeable plastic torsion rods. The spring constants of such rods are determined by their length and diameter. Rods could be preset so that any desired amount of torsion was present

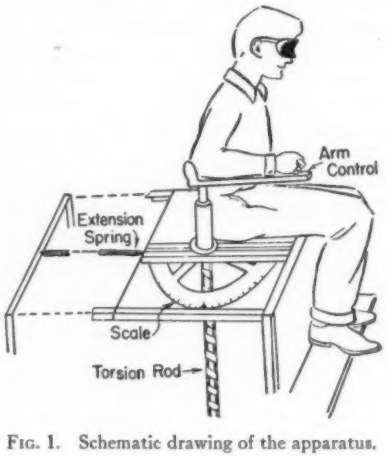

Fig. 1. Schematic drawing of the apparatus.

TABLE 1

CONTROL LOADINGS USED FOR EACH AMPLITUDE OF MOTION

Torque at Starting Position	Torque at Terminal Position			
	0 in.-lbs.	50 in.-lbs.	100 in.-lbs.	200 in.-lbs.
0 in.-lbs.	x	x	x	x
25 in.-lbs.		x		
50 in.-lbs.		x	x	
100 in.-lbs.			x	x
200 in.-lbs.				x

Note.—For the 50–50, 100–100, and 200–200 conditions the initial torques actually were slightly smaller than, but not discriminable from, the terminal torques.

when the control was in its starting position and resting against a stop.

Another type of spring loading was used in which the level of spring tension remained relatively constant for all positions of the control, i.e., the change in the length of the spring for various positions of the control was such a small fraction of the total length of the spring that the associated changes in force were not discriminable. This type of loading was achieved by substituting 15-ft. long coil springs for the torsion rods.

The arm control could be rotated a maximum of 90° clockwise in the horizontal plane. Movements could be recorded in three different ways. For the first method of recording, a plastic rider was displaced along a degree scale as the control was rotated clockwise, and remained in position when the control was returned to its starting point. The maximum excursion of the control was read off the scale and the rider returned to its original position for the next movement. For the second method of recording, a pointer was attached so that it moved over the degree scale during the course of the movement. Angular position was read to the nearest .5° as *S* attempted to maintain the control in a constant position. The third scoring method utilized graphic records obtained by means of a Sanborn recorder which responded to voltage changes across a potentiometer attached to the control shaft.

The length of the arm control was adjusted to the arm length of each *S*. However, the forces exerted upon the control by *S*s can be determined approximately by dividing all torque values by 12, the average length.

Procedure

Each *S* was tested on each of 30 different conditions under two sets of instructions. Under the set of instructions employed in Part 1, *S*s made a single rapid movement and returned the control immediately to its starting position. The maximum excursion of the control was recorded in this case. Under the instructions employed in Part 2, *S*s could make corrective movements and the position of the control was recorded at the point at which *S* indicated he had completed these secondary adjustments.

Part 1.—Seven male students at The Ohio State University, paid by the hour, served as *S*s.

The *S*s were seated, the control handle was adjusted according to their arm length, and blackout goggles were placed over their eyes.

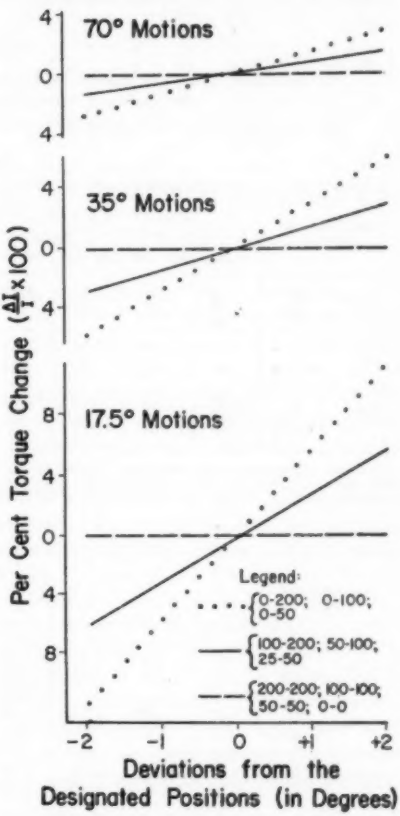

FIG. 2. Relative torque change as a function of amplitude deviation.

They were instructed to move the control rapidly by the indicated amplitude, to return it immediately to its original position without corrective movements, and were told that the maximum excursion of each movement would be scored. An electric timer gave an auditory signal at 10-sec. intervals; *S*s were instructed to initiate a new movement each time they heard the auditory signal. After each response *S* was told the amount by which he had overshot or undershot to the nearest .5°. Three amplitudes (17.5°, 35°, and 70°) of movement of the lower arm in the horizontal plane with the elbow fixed were studied, and ten combinations of initial and final torque were studied with each amplitude of motion. The ten torque conditions are indicated in Table 1. The relative torque change as a function of deviations from the designated amplitude are shown graphically in Fig. 2 for all 30 conditions. These 30 combinations of torque and amplitude were given to each *S* in a different sequence selected from a table of random numbers. Each *S* served for 55 consecutive trials on each condition; a rest period of 1 min. was given after each 11th trial. A rest period of at least 1 hr. was given at the end of each condition.

Part 2.—Six male *S*s served in Part 2. Three of these had served previously in Part 1.

The 30 experimental conditions were the same as in Part 1, but they were administered to each *S* in a new random sequence. The *S*s were now instructed to make corrective movements if they thought they had overshot or undershot the correct distance, and to say "now" when they thought they had achieved the correct position. This designated position was recorded to the nearest .5°.

TABLE 2

AVERAGE ABSOLUTE ERROR IN DEGREES PER MOVEMENT FOR MOVEMENTS OF VARYING AMPLITUDE AND TORQUE CHANGES

The error for each *S* was computed separately and the data were then pooled for all *S*s

Part 1

Amplitude of Motion		Torque Change (in.-lbs.)									
		0–200	100–200	200–200	0–100	50–100	100–100	0–50	25–50	50–50	0–0
17.5°	*M*	.67	.92	1.33	1.14	1.22	1.21	1.08	1.14	1.29	1.34
	SD	.29	.35	.48	.67	.48	.33	.45	.30	.48	.41
35°	*M*	1.11	1.73	2.22	1.61	1.44	1.59	1.90	2.29	2.01	1.88
	SD	.33	.47	.79	.44	.57	.52	.80	.73	.85	.51
70°	*M*	1.87	2.11	2.32	2.53	2.32	2.52	2.60	1.78	2.24	2.94
	SD	.54	.84	.64	.93	.83	.84	.52	.62	.61	1.10

Note.—Using an estimate of error pooled for all 10 conditions for any one extent of motion the difference between means necessary for t to be significant at $p \leq .01$ is .41 for 17.5°, .62 for 35°, and 1.05 for 70°. Corresponding values for $p \leq .05$ are .31 for 17.5°, .47 for 35°, and .79 for 70°.

Part 2

Amplitude of Motion		Torque Change (in.-lbs.)									
		0–200	100–200	200–200	0–100	50–100	100–100	0–50	25–50	50–50	0–0
17.5°	*M*	.61	.84	1.02	.71	.89	.92	.82	.98	.50	1.12
	SD	.10	.26	.16	.18	.06	.08	.21	.30	.23	.37
35°	*M*	1.05	1.31	1.75	1.14	1.27	1.41	1.23	1.24	1.46	1.54
	SD	.11	.26	.34	.14	.34	.23	.14	.28	.16	.30
70°	*M*	1.39	1.48	1.84	1.56	1.84	1.43	1.91	1.54	1.41	2.30
	SD	.30	.53	.40	.27	.47	.35	.39	.29	.20	.78

Note.—Using an estimate of error pooled for all 10 conditions for any one extent of motion the difference between means necessary for t to be significant at $p \leq .01$ is .34 for 17.5°, .34 for 35°, and .46 for 70°. Corresponding values for $p \leq .05$ are .25 for 17.5°, .25 for 35°, and .61 for 70°.

Results and Discussion

Average errors were computed separately for successive blocks of 11 responses in all experimental conditions and learning curves plotted. From inspection it appeared that nearly all of the learning occurred during the first 22 responses on any condition. Of the remaining responses in each condition, all except the first response immediately following each of the three rest periods (a total of 30 responses) were used in the analyses that follow.

The average error (regardless of sign) in degrees per movement is shown in Table 2 for all conditions of Part 1 and Part 2. These data are pooled for all *S*s and the error is computed with respect to the reference positions (17.5°, 35°, and 70°), not with respect to *S*'s own mean.

The data in Table 2 for 27 of the 30 conditions were analyzed separately for Part 1 and Part 2 by means of a four-factor analysis of variance design with repeated measures on the same *S*s. Variance among *S*s was one of the four factors; the others were (*a*) amplitude of movement, (*b*) torque at termination of movement, and (*c*) ratio of initial to terminal torque. The conditions of zero spring loading were not included since they represented a special case; both analyses thus included 27 conditions (3 amplitudes × 3 terminal torques × 3 ratios of initial to terminal torque). The results of the two analyses are shown in Tables 3 and 4, respectively.

With terminal torque held constant, average absolute error increases with amplitude of motion, and this increase is highly significant for both analyses. However, the ratio of average error to amplitude of motion decreases with amplitude. The relative error is

TABLE 3

Analysis of Variance for Uncorrected Movements, Part 1

Source	*df*	Mean Square	Error Term	*F*
1. Amplitude (A)	2	20.6950	5	24.40**
2. Terminal torque (T)	2	.8560	5	1.01
3. Ratio of initial to terminal torque (R)	2	1.0512	12	1.96
4. Subjects (*S*)	6	4.2860	12	8.00**
5. A × T	4	.8481	15	5.55**
6. A × R	4	.4350	12	.81
7. A × S	12	.6005	12	1.12
8. T × R	4	1.0013	15	6.55**
9. T × *S*	12	.0544	15	.36
10. R × *S*	12	.4726	12	.88
11. A × T × R	8	.2706	15	1.77
12. A × R × *S*	24	.5356	15	3.51**
13. A × T × *S*	24	.2360	15	1.54
14. T × R × *S*	24	.2409	15	1.58
15. A × T × R × *S*	28	.1528		

Note.—The significance of all terms is tested against the next higher order term which is significant and which contains that term.
** Significant at $p \leq .01$.

TABLE 4

Analysis of Variance for Corrected Movements, Part 2

Source	*df*	Mean Square	Error Term	*F*
1. Amplitude (A)	2	7.6724	Pooled 6 and 7	40.78**
2. Terminal torque (T)	2	.0183	8	.38
3. Ratio of initial to terminal torque (R)	2	.4840	Pooled 6 and 8	1.29
4. Subjects (*S*)	5	.5215	7	3.40*
5. A × T	4	.0322	15	.44
6. A × R	4	.2750	15	3.73*
7. A × *S*	10	.1534	15	2.08
8. T × R	4	.4762	15	6.45**
9. T × *S*	10	.0597	15	.81
10. R × *S*	10	.1295	15	1.75
11. A × T × R	8	.0851	15	1.15
12. A × R × *S*	20	.0693	15	.94
13. A × T × *S*	20	.0416	15	.56
14. T × R × *S*	20	.0586	15	.79
15. A × T × R × *S*	40	.0738		

Note.—The significance of all terms is tested against the next higher order term which is significant and which contains that term.
* Significant at $p \leq .05$.
** Significant at $p \leq .01$.

6.48%, 5.05%, and 3.31% for 17.5°, 35°, and 70° of movement, respectively, in Part 1, and 5.03%, 3.38%, and 2.39%, respectively, in Part 2.

The poorest mean accuracy for movements of any given amplitude was consistently found for the conditions of zero spring loading for all Ss. Of the 54 conditions in which some spring loading was used, only three conditions gave higher average errors than the movements of comparable amplitude involving no spring loading, and these three instances are probably accounted for by chance variations in performance.

The effects of absolute and relative torque change per degree of movement are similar for both methods of executing the movement, and are in the direction predicted by the original hypotheses. High ratios of terminal to initial torque generally resulted in small terminal position errors for the 17.5° and 35° amplitudes of motion. In Part 2 the effect of torque ratio was not significant when tested against the interaction of torque ratio with movement amplitude, but the latter interaction is significant at the .02 level of confidence when tested against the third-order interaction term. This interaction apparently arises from the fact that torque change is relatively more beneficial with small than with large movement amplitudes. In the case of uncorrected movements (Part 1) the differences are all in the same direction as in Part 2, and the interaction of amplitude with torque ratio is significant when tested against the third-order interaction term. However, *S* variability for this interaction effect is significant as indicated by the triple interaction term of Ss × amplitude × torque ratio, and the effects of torque ratio and torque ratio × amplitude interaction are not significant when tested against this triple interaction term.

High ratios of terminal to initial torque are more effective in reducing errors when terminal torques are larger, i.e., when the absolute torque change per degree of movement is greater. This is indicated by the interaction of torque ratio × terminal torque which is significant at the .01 level of confidence for both parts of the experiment.

A graphic presentation of the effects of torque changes upon accuracy of positioning the control is shown in Fig. 3. In plotting this graph, data for conditions involving equal relations between position error and percentage of torque change, $\frac{\Delta I}{I(\Delta D)}$, have been combined from both parts of the experiment regardless of the torque at the terminal position. Thus for each amplitude of movement the 50–50, 100–100, and 200–200

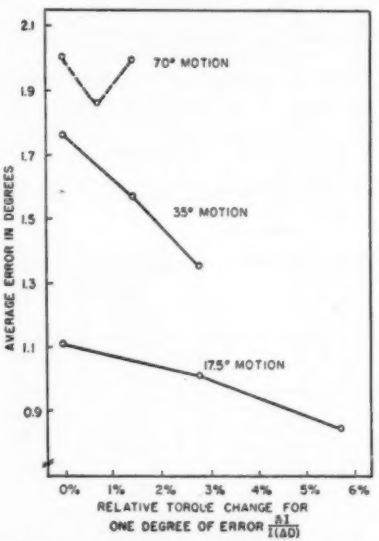

FIG. 3. Average error as a function of torque change.

TABLE 5

TREND ANALYSIS FOR THE EFFECT OF RELATIVE TORQUE CHANGE UPON POSITIONING ERROR

Source of Variance	*df*	17.5° Motions		35° Motions		70° Motions	
		Mean Square	*F*	Mean Square	*F*	Mean Square	*F*
Between means for individuals	11	2.56	5.82*	4.68	7.43*	2.87	0.72
Between slopes for individuals	11	0.12	0.27	0.24	0.38	0.40	0.10
Between group means	1	5.81	13.20*	17.53	27.83*	37.29	9.42
Between group slopes	1	0.10	0.23	0.01	0.02	0.11	0.03
Over-all slope	1	4.36	9.91*	9.21	14.62*	0.01	0.00
Over-all deviations from linearity	1	0.03	0.07	0.02	0.03	1.45	0.37
Group deviations from estimation	1	0.19	0.43	0.42	0.67	1.83	0.46
Individual deviations from estimation	11	0.44		0.63		3.96	

* Significant at the $p \leq .01$ level of confidence when tested against respective individual deviations from estimation.

torque conditions are combined, the 25–50, 50–100, and 100–200 conditions are combined, and the 0–50, 0–100, and 0–200 conditions are combined just as in Fig. 2. Each point on the graph thus represents a weighted average of the data collected under six conditions, three taken from each of the two parts of the experiment.

The significance of the slopes shown in Fig. 3 was tested by means of an Alexander trend analysis (1). The results of this analysis are shown in Table 5. The over-all slope is significant at $p \leq .01$ for 17.5° and 35° motions, but not for 70° motions. None of the slopes show reliable departures from linearity, and the slopes do not differ significantly for corrected and uncorrected movements of the same magnitude. In view of the latter result, the pooling of the data for corrected and uncorrected movements appears to be justified in examining slope effects.

The graph illustrates the combined effect of relative and absolute torque changes, since the conditions involving larger relative torque changes also involve larger absolute torque changes. It should be noted, however, that mean terminal torque is the same for all of the nine points plotted in Fig. 3. Independent plots of the effects of relative and absolute torque change per degree of movement would only be warranted if data were collected under a large number of conditions and on a larger sample of *S*s. The treatment of the data shown in Fig. 3 indicates clearly the beneficial effects of torque changes upon the accuracy of the 17.5° and particularly on the 35° movements. For the 70° movements no such effect is apparent.

Constant errors were analyzed in the same manner as variable errors. All of the mean errors were quite small in magnitude and none of the *F* ratios were significant at the .05 level of confidence. It appears that *S*s are able to reduce their constant errors to a minimum when immediate knowledge of results is given and when a large number of responses are made in succession on the same condition.

The results lend confirmation to the initial hypotheses. It appears that torque changes furnish very useful cues for the type of movements

studied, provided the relative change of torque per degree of error in position is not less than approximately 3%, and the absolute torque change equals approximately 6 in.-lbs. per degree of movement. Smaller relative changes might be useful with larger absolute changes and vice versa.

It is concluded that three variables, (*a*) amplitude, (*b*) terminal torque, and (*c*) relative torque change per unit of amplitude change, all influence the accuracy of positioning motions. Several of the interactions among these variables also appear to be important.

Summary

This study investigated accuracy of positioning a horizontal arm control as a function of changes in the torque-displacement relation of the control. It was found that positioning errors were smallest when the ratio of relative torque change to displacement is largest, particularly if the absolute change of torque with displacement is also large.

It is concluded that control forces opposing a movement can provide useful cues in learning to execute different amplitudes of movement, and that their usefulness depends upon the relative and absolute torque change per unit of control motion.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Bahrick, H. P., Fitts, P. M., & Schneider, R. The reproduction of simple movements as a function of factors influencing proprioceptive feedback. *J. exp. Psychol.*, 1955, **49**, 445–454.
3. Jenkins, W. O. The discrimination and reproduction of motor adjustment with various types of aircraft controls. *Amer. J. Psychol.*, 1947, **60**, 397–406.
4. Orlansky, J. Psychological aspects of stick and rudder controls in aircraft. *Aero. Engng Rev.*, 1949, **8**, 1–10.
5. Weiss, B. The role of proprioceptive feedback in positioning responses. *J. exp. Psychol.*, 1954, **47**, 215–224.

(Received for early publication March 3, 1955)

Journal of Experimental Psychology
Vol. 49, No. 6, 1955

REPRODUCTION OF SIMPLE MOVEMENTS AS A FUNCTION OF FACTORS INFLUENCING PROPRIOCEPTIVE FEEDBACK [1]

HARRY P. BAHRICK,

Ohio Wesleyan University

PAUL M. FITTS AND RONALD SCHNEIDER [2]

The Ohio State University

Hypotheses regarding the role of proprioception in skilled motor performance are not easily accessible to experimental verification because of the difficulty of controlling the proximal stimulus. Apart from techniques employing surgery or drugs the most promising approach to this topic appears to be that of varying the type and degree of resistance to movement offered by a control which *S* uses in the performance of a motor task. The nature of this resistance is a function of four physical properties of the control—the mass, viscosity, elasticity, and degree of coulomb friction associated with its movement—and these physical parameters can all be controlled independently by *E*.

Proprioceptive information is used by *S* in at least two ways. If movements continue for more than the duration of a reaction time, proprioceptive feedback arising during the initial phases of a motor response may provide knowledge useful in regulating later phases of the same movement sequence. This use of proprioceptive information therefore aids *S* in the regulation of continuing activity. The other use of proprioceptive information can apply to movements of any duration. It refers to the facilitating effect of proprioception on the learning of a particular motion, which enables *S* to organize and initiate a sequence of motor activity with more appropriate force and timing as practice continues.

With respect to both ways in which proprioceptive information may be used, it is hypothesized that, within limits, an increase in the resistance to movement offered by a control should facilitate performance by increasing the discriminability of different patterns of control movement. This effect is achieved at the expense of increasing the range of forces required of a movement, since control loading is in all instances an additive effect. The direction of the hypothesized effect is, of course, contrary to what would be predicted from a thesis of minimal effort. It is assumed that an increase in the range of forces required in a graded or patterned activity increases the number of absolutely discriminable response categories by increasing the range of intensities of proprioceptive feedback which can be associated with the execution of different response patterns.

Several studies (**2, 4, 5, 6, 7, 8, 9**) have investigated accuracy of movement as a function of the physical characteristics of controls, although

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-10528 with the Ohio State University Research Foundation, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, publication, use, and disposal in whole or in part by or for the United States Government.

[2] Now at the Aero Medical Laboratory, Wright-Patterson Air Force Base.

the tracking study by Howland and Noble (6) is the only one in which the effect of three parameters (elasticity, damping, and mass) of a control were investigated concurrently.

The purpose of the present study was to determine the effect of three different kinds of resistance to movement offered by a control on the ability of *S*s to replicate simple movements with the control. The *S*s were instructed to make simple triangular and circular movements in time to a metronome. Their movements were scored for spatial and temporal accuracy of reproductions based on proprioceptive memory.

The following hypotheses were tested: (*a*) accuracy in moving a control to a prescribed position is increased by increasing the elasticity of a control; (*b*) accuracy in maintaining a prescribed rate of motion is increased by adding viscous damping or mass to a control. More generally, these hypotheses propose that elasticity added to a control should lead to improved spatial accuracy of movements, while damping and mass added to a control should lead to improved temporal accuracy of movement.

Each of these hypotheses is suggested by a corresponding physical relation (3, p. 1317). Thus (*a*) the relation between the torque applied to a control and the resulting positional displacement is a function of the elasticity of the control, (*b*) the relation between applied torque and the resulting rate of motion is a function of viscous damping, and (*c*) the relation between applied torque and resulting acceleration is a function of the mass (or the moment of inertia) of the control. Briefly, the relations among the physical parameters of the control, the intensity of stimulation arising from proprioceptive feedback, and the predicted effects on movement accuracy can be illustrated by considering the nature of the control forces that oppose a time-varying torque L_t applied by *S* to the handle of a control such as that used by Howland and Noble (6). Disregarding for the moment the mass, viscosity, elasticity, and coulomb friction of the arm, the time-varying torque (L_t) applied by *S* is opposed by the mass, viscosity, and elasticity of the control in the manner indicated by a system equation of the form

$$L_t = K\theta + B\frac{d\theta}{dt} + J\frac{d^2\theta}{dt^2}, \qquad (1)$$

where K is the constant of elasticity, B is the constant of viscosity, J is the moment of inertia of the control, and θ is the angular displacement of the control handle with respect to some neutral (spring-centered) position. The dynamic relations expressed by this equation can be illustrated by considering the case in which a constant torque L_c is applied to a control which possesses no spring loading. Since in this case $K = 0$, the torque applied to the control at the point of contact with *S*'s limb would be opposed by equal forces proportional to (angular) rate and acceleration of motion, but not by any forces proportional to angular displacement. Here *S* would derive no additional proprioceptive cues which could aid him in learning to associate a particular torque with a particular displacement of the control. If the control possesses a large K factor, however, a constant applied torque will cause it to come to rest with a specific terminal displacement, a position for which $L_c = K\theta$, and this terminal position can become associated with the proprioceptive stimulation arising during the response made with this particular applied torque. Thus, the K factor of the system provides a basis on which *S* may be able to learn to discriminate among various terminal control positions.

The argument for the effect of damping and mass on learning to use proprioceptive cues is similar to that for elasticity. If the control possesses no damping characteristic, the constant torque applied to the control handle would be opposed only by torques pro-

portional to displacement and acceleration. Thus, this type of control would not provide cues that could become associated with a particular rate of arm motion. However, if we assume a large damping constant but no elasticity and minimum inertia, any constant torque applied to the control would result in a corresponding constant rate of motion, so that $L_t = B d\theta/dt$. The conditions for learning and performance are thus maximized with respect to rate of motion when we use a large damping constant since rate is now almost directly proportional to torque and since we have also increased the range of discriminable forces which the arm is required to produce. All systems will, of course, possess some inertia, and therefore the application of a constant torque will not immediately result in a constant rate of angular motion. If the system also possesses some K, the constant torque applied by S will result in a gradually decelerating, rather than a constant, rate of motion and these variations will tend to interfere with the association of a particular rate with a particular applied torque.

Finally, the presence of inertia in the system will enable S to associate various torques with corresponding accelerations provided that B and K in the system are small. The presence of substantial damping or elasticity constants in the system would limit this possibility.

The preceding discussion is somewhat hypothetical because previous experiments (**2, 8**) have shown that it is difficult for S to exert a constant torque against a movable control particularly at higher velocities, probably due in part to the fact that he must also accelerate the mass of his own arm. Although this practical limitation does not necessarily limit the value of the theoretical analysis, it should be kept in mind, nevertheless, that we are considering the effect of adding elasticity, damping, and mass to a control, and that these additional loads augment the physical characteristics of the limb itself. If the former effects can be predicted, then it could be postulated further that the effects of elasticity, damping, and mass internal to the arm itself are similar to those arising from the addition of such constants to an external control.

Method

Apparatus.—A schematic drawing of the joystick control is shown in Fig. 1. The maximum radius of control movement was 9 in. Four identical springs were attached at right angles to each other at point S on the control stick. With this type of spring attachment an equal torque in any direction will displace the control an equal distance from its center position. Twelve sets of springs, differing in spring constant, were used. Damping was produced by attaching hollow aluminum drums of three diameters (3 in., 5½ in., and 8½ in.) to the lower end of the control and by filling the tank with water. Holes were pierced in the surface of the drums in order to minimize turbulence and backflow in the water. Inertia was controlled

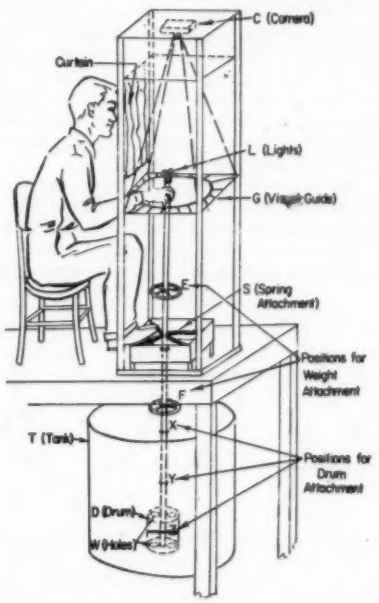

Fig. 1. Schematic drawing of the apparatus.

by adding equal amounts of lead weights to the control at positions E and F (Fig. 1) equidistant on either side of the fulcrum.

Two easily removable visual guides (G in Fig. 1), one triangular and the other circular, were used to indicate to *S* the required patterns of movement. They were made of rubber strands fastened by thread to a larger wooden frame. The control knob could be moved directly along the periphery of these guides. For the circular guide the rubber strand was attached to the frame by means of 32 strings, so that the guide was actually an equilateral polygon of 32 sides. The coefficient of elasticity of the rubber guides was so small that they did not provide any substantial kinesthetic stimulation when touched by the control, and therefore provided visual guidance only.

An electric metronome provided an auditory signal which was used to pace *S*'s speed of movement during practice trials. A black curtain could be dropped between *S*'s face and the control to aid in minimizing visual cues.

Two small lights were mounted side by side on top of the joystick (L in Fig. 1). One light remained on continuously while the other was flashed intermittently at a rate of 5.6 times per second.

A 16-mm. camera was mounted 27 in. directly above the center position of the joystick (C in Fig. 1). It had a wide angle lens and a single-frame time exposure attachment. Each triangular or circular movement of the control was recorded on a single frame of 16-mm. film as two lines traced by the lights, one continuous and the other intermittent. Tracings of typical records for circular and triangular control movements are shown in Fig. 2.

Procedure.—A preliminary study was conducted to determine values of *K*, *B*, and *J* that were equated with respect to subjective judgments regarding the amount of "effort" involved in executing circular and triangular movements under different conditions of control loading. A variation of the constant stimulus method was employed with nine groups of 10 *S*s each. Separate values of the control constants were determined for the two types of movements, circular and triangular. A more detailed account of the procedure in the preliminary study is given elsewhere.[3]

FIG. 2. Typical record of a triangular and a circular motion. The continuous line indicates the path of the tip of the joystick control; the distance between dots is the distance travelled in .18 sec.

In the experiment proper the accuracy of movement reproduction was studied as a function of the type and degree of control loading. A triangular movement was studied under nine conditions—three values of *K*, three values of *B*, and three conditions of *J*. The three values of *K* were selected arbitrarily. These were 40, 80, and 160 in.-oz. per degree of displacement of the control from the center position. Each of the values of *K* was then matched to a corresponding value of *B* and *J* on the basis of the data from the preliminary psychophysical study. Only one type of physical constant was employed at a time. A circular movement was studied in a similar manner, under three values of *K* and three corresponding values of *B*. The conditions of added mass were not included in this part of the experiment because the preliminary study indicated that in the case of a circular movement the amount of mass required to produce subjective equality of effort, as compared to *K* and *B* effects, exceeded the capability of the apparatus at the speeds employed.

Twenty male and 20 female volunteer novice undergraduate students at Ohio Wesleyan University served as *S*s under all 15 experimental conditions. Sequential effects were controlled by assigning a different sequence of conditions, obtained from a table of random numbers, to each *S*.

Immediately preceding each series of test trials each *S* was given three practice trials in making the triangular or circular movement, using the visual guides described earlier. The value of spring, damping, or mass to be used during the test trials was present during practice trials. All practice trials on the triangular movements were paced by the metronome at a rate of 176 beats per minute, and *S*s were instructed that each side of the triangle should be completed in two beats. The metronome was set at 138 beats per minute for the circular movements and *S*s were instructed to traverse each quadrant in one beat.

Following each set of practice trials the metronome and guide were removed, the curtain

[3] A more detailed description of the psychophysics procedure has been deposited with the American Documentation Institute. Order Document No. 4518, remitting $1.25 for microfilm or $1.25 for photocopies.

TABLE 1

SCORES DERIVED FROM ANALYSES OF PHOTOGRAPHIC RECORDS OF MOVEMENT

Score	Measure Taken from Each Motion	Scores for the Individual
	Triangular Motions	
1a, 1b, 1c	Maximum speed on sides a, b, and c, respectively	Variability of measures a, b, and c, respectively, over seven trials
2	Same as for 1a, 1b, 1c	Variability among a, b, and c within a single trial averaged for seven trials
3	Total movement time for a trial	Mean of seven successive trials
4	Same as for (3)	Variability among seven successive trials
5a, 5b, 5c	Spatial position of sides a, b, and c, respectively	Median position error for seven successive trials for a, b, and c, respectively
6a, 6b, 6c	Same as for 5a, 5b, 5c	Variability among seven successive trials for a, b, and c, respectively
	Circular Motions	
7	Maximum and minimum speed exclusive of initial and final acceleration	Difference between maximum and minimum speed within a given trial averaged for seven trials
8	Same as for (7)	Difference between maximum and minimum speed at any time over seven successive trials
9	Total movement time for a trial	Mean of seven successive trials
10	Same as for (9)	Variability among seven successive trials
11a, 11b, 11c	Radius at 90°, 180°, and 270° (movement starts at 0°), respectively	Median radial error for seven successive reproductions for a, b, and c, respectively
12a, 12b, 12c, 12d	Radius at 45°, 135°, 225°, and 315°, respectively	Radial error variability over seven successive trials for a, b, c, and d.

was placed between *S*'s face and the control, and the room was darkened. The *S* was instructed to duplicate the movement just practiced, paying attention both to spatial pattern and to timing. The *E* returned the control to the starting position after each trial and gave no knowledge of results. Seven trials, exclusive of the practice trials, were given for each of the 15 conditions. There was a 5-sec. interval between trials. During each test trial the movement was photographed by the 16-mm. motion picture camera, using a single frame time exposure for each trial. A 30-sec. rest interval was given between conditions.

Scores obtained from photographic records.— Each record was projected on a screen and 23 measures of temporal and spatial accuracy, including measures of mean and median error and variability, were computed. These measures are listed in Table 1.

RESULTS

Means and *SD*s for each type of score and each control condition are reported in Table 2 for the triangular motions and in Table 3 for the circular motions. A separate analysis of variance for repeated measures on the same *S* was performed for each of the 12 scores computed for the triangular motions. In each analysis the significance of variation among the three parameters of elasticity, damping, and mass was first tested. Parameter variance was evaluated against the first order parameter × *S* interaction with 2 and 2 × 39 *df*. Next, the significance of the variation among the three spring conditions, among the three damping conditions, and among the three conditions of mass was tested. For these tests the variance among the three conditions of a parameter was evaluated against the respective conditions × *S* interaction

TABLE 2

MEANS AND *SD*s FOR ALL SCORES, AND *p* VALUES FOR TESTS OF SIGNIFICANCE OF DIFFERENCES AMONG AND WITHIN PARAMETERS

Triangular Movements

Control Loading Condition	Statistic	Score											
		Temporal Measures						Spatial Measures					
		1a	1b	1c	2	3	4	5a	5b	5c	6a	6b	6c
Light spring	*M*	2.02	2.66	2.00	2.61	12.39	1.70	1.19	1.41	1.30	2.98	3.61	3.27
	SD	.68	.96	.65	.80	1.08	.68	.74	1.11	.74	.91	1.21	1.33
Medium spring	*M*	1.98	2.21	2.22	2.77	12.28	1.58	1.26	1.88	1.50	2.89	3.70	3.06
	SD	.70	.74	.82	.94	.91	.74	.74	1.29	.91	.80	1.04	.89
Heavy spring	*M*	2.42	2.29	2.82	3.34	12.09	1.75	.99	1.86	1.18	3.41	3.72	3.06
	SD	.81	.98	1.02	.94	1.22	.73	.80	1.38	.90	1.19	1.02	1.18
Variance within *K*	*p*	<.02	—	<.01	<.01	—	—	—	—	—	—	—	—
Light damping	*M*	1.88	2.02	1.76	2.58	13.15	1.72	1.52	1.86	1.25	2.96	3.92	3.11
	SD	.92	.64	.48	.71	1.32	.92	1.00	1.45	.89	1.43	1.26	.93
Medium damping	*M*	1.50	1.71	1.63	2.14	13.72	1.85	1.49	1.60	1.01	3.06	3.73	2.99
	SD	.52	.57	.67	.65	1.20	.73	1.10	1.42	.92	1.05	1.23	.96
Heavy damping	*M*	1.67	1.44	1.56	2.07	14.56	1.98	1.62	2.22	1.21	3.29	3.96	3.06
	SD	.80	.63	.50	.80	1.26	.85	1.06	1.39	.94	1.12	1.12	1.18
Variance within *B*	*p*	—	<.01	—	<.01	<.01	—	—	<.05	—	—	—	—
Light mass	*M*	1.94	1.92	1.95	2.48	12.96	1.75	1.46	2.12	1.16	2.87	4.25	2.99
	SD	.75	.76	.76	.76	1.21	.97	1.01	1.47	.87	.89	1.22	.90
Medium mass	*M*	1.68	1.90	1.74	2.26	13.54	1.90	1.48	2.06	1.32	3.22	4.14	2.91
	SD	.71	.63	.79	.64	1.33	1.02	1.12	1.40	1.04	1.14	1.59	.98
Heavy mass	*M*	1.52	1.47	1.65	2.03	14.22	1.85	1.56	1.71	1.26	3.01	4.39	3.10
	SD	.64	.60	.68	.72	1.35	.88	1.02	1.30	.83	1.01	1.41	.63
Variance within *J*	*p*	<.02	<.01	—	<.01	<.01	—	—	—	—	—	—	—
Variance among parameters	*p*	<.01	<.01	<.01	<.01	—	—	<.01	—	—	—	—	—

TABLE 3

MEANS AND STANDARD DEVIATIONS OF ALL SCORES FOR ALL CONDITIONS AND p VALUES FOR TESTS OF THE SIGNIFICANCE OF DIFFERENCES AMONG AND WITHIN PARAMETERS OF CONTROL LOADINGS

Circular Movements

Control Loading Condition	Statistic	Score										
		Temporal Measures				Spatial Measures						
		7	8	9	10	11a	11b	11c	12a	12b	12c	12d
Light spring	*M*	3.34	4.15	12.32	2.02	2.02	1.90	2.38	3.78	3.72	4.59	3.98
	SD	.92	1.10	1.23	.82	1.50	1.52	1.54	1.20	1.13	1.36	1.23
Medium spring	*M*	3.84	4.74	12.26	2.02	1.95	1.92	2.51	3.92	3.96	4.74	4.52
	SD	1.13	1.30	1.55	1.17	1.22	1.74	1.84	1.17	1.31	1.52	1.62
Heavy spring	*M*	4.34	5.17	11.95	2.02	1.74	1.78	2.62	4.11	3.71	4.46	3.98
	SD	1.00	1.01	1.26	.76	1.19	1.32	1.48	1.29	1.01	1.38	.98
Variance within *K*	*p*	<.01	<.01	<.01	—	—	—	—	—	—	—	—
Light damping	*M*	3.32	4.14	12.62	2.02	1.95	2.30	2.67	3.75	4.12	4.84	3.82
	SD	.96	1.06	1.30	.94	1.45	1.92	1.94	1.44	1.29	1.63	1.52
Medium damping	*M*	2.99	3.48	13.32	1.95	2.15	2.78	2.48	3.49	3.96	4.90	3.85
	SD	.75	.76	1.55	1.22	1.48	2.27	1.94	1.12	1.26	1.40	1.12
Heavy damping	*M*	2.87	3.38	13.85	1.90	2.16	2.06	2.87	3.82	3.94	4.38	3.62
	SD	.70	.81	1.47	.83	1.47	1.55	1.82	1.20	.99	1.21	.99
Variance within *B*	*p*	<.01	<.01	<.01	—	—	—	—	—	—	—	—
Variance among parameters	*p*	<.01	<.01	<.01	—	—	<.05	—	—	—	—	<.05

with 2 and 2 × 39 *df*. In order to conserve space only the *p* values corresponding to the obtained *F* ratios are given in Table 2 for each of the 13 analyses.

Comparable analyses were performed for the 11 types of scores obtained from the circular motions. Significance of the variation between the parameters of elasticity and damping was first evaluated against the parameter × *S* interaction with 1 and 1 × 39 *df*. Variation among the three conditions of elasticity and among the three damping conditions was then evaluated separately against the respective *S* × condition interaction variance with 2 and 2 × 39 *df*. The *p* values computed for the obtained *F* ratios are reported in Table 3.

Triangular Movements

Temporal accuracy.—Uniformity of peak speed over seven trials in each leg of the triangle was improved significantly by the damping conditions and by the conditions of increased mass in the control, while the spring-loaded conditions resulted in a significant decrease in response uniformity. The effect of changing the level of *K*, *B*, and *J* on peak speed uniformity was in the same respective directions, was consistent for all three sides of the triangle, and was significant at least at the .02 level of confidence for seven of the nine comparisons. However, no significant differences were found over seven successive trials in the variability of the time taken for the entire movement. An increase in the magnitude of either mass or damping tended to increase the total time for all movements, while increasing spring tension decreased total movement time by a small, but nonsignificant amount.

Spatial accuracy.—In general the different loading conditions resulted in nonsignificant differences in spatial accuracy. Significant differences between control parameters were obtained only for the first leg of the triangle with largest value of *K* leading to the smallest spatial error.

Circular Movements

Temporal accuracy.—Variability of rate within a single movement was significantly smaller with damped conditions than with spring-loaded conditions. The greater the value of *B* the smaller the variability of rate; the greater the value of *K* the greater the variability of rate. These effects of increasing *B* and *K*, which are in accordance with the original hypotheses, are significant at the .01 level of confidence. Significant differences were also obtained for measures of variability in rate among successive reproductions. However, the variability in total movement time was not significantly different between *B* and *K* conditions, and was not significantly affected by variation in the level of *B* and *K*. Total movement time was increased by a significant amount under the damping conditions; it also increased significantly for the larger values of *B*, and decreased by a small amount for the larger values of *K*.

Spatial accuracy.—As was the case for triangular movements, the different loading conditions resulted in general in nonsignificant differences in spatial accuracy of replication. Only in the second quadrant of the circle did spring-loading conditions result in a significantly smaller median radial error than the damping conditions.

Discussion

Of the 23 separate scores computed for the two types of movement 10 were

measures of temporal characteristics and 13 were measures of spatial characteristics. The timing measures consistently revealed statistically significant differences (*a*) among the three control parameters and (*b*) within the three levels of each parameter. Significant effects in the *variability* of temporal aspects appeared within the different segments of movements, however, and not in the over-all time to complete the movements. None of the measures of spatial characteristics of movements revealed significant variability within levels of any of the three parameters, and only one measure revealed a difference among the three parameters that was significant at the .01 level of confidence.

The following conclusions are drawn regarding the significant effects on the time measures resulting from the addition of mass, damping, and stiffness to a control: (*a*) Under conditions of damping a particular movement tends to be made at a significantly more uniform rate, and the maximum rate attained from one movement to the next is also more uniform; (*b*) movements made with a control having added mass are made at more uniform speed, and the peak speed tends also to be more uniform from one movement to the next; (*c*) the addition of spring stiffness impairs the maintenance of uniform speed, and increases variability in maximum attained speed. In general, these effects were uniformly progressive for each of the three levels within each parameter, maximum mass and maximum viscosity being maximally beneficial and maximum spring constant being maximally harmful.

The observed effects of damping and mass are in agreement with the predictions. The differential effects produced by varying the amounts of damping and of mass further substantiate the interpretation that damping and mass facilitate rate and acceleration discriminations, respectively, whereas elasticity makes such discriminations more difficult. These relative effects of varying amounts of mass, spring, and damping are discussed in detail in another paper (now in preparation).

It is possible, however, to interpret the effects of damping and mass without hypothesizing that the effect is related to the use of proprioceptive information. The addition of damping to a control requires *S* to exert a greater force in order to move the control at any prescribed rate. Thus, increased discriminability is achieved by requiring the motor system to operate over a greater range of forces. The general reduction of average speed for all movements with increased damping or mass is probably due to this physical effect, and it is possible that the decreased variability in the temporal aspects of movements under these conditions could be determined by the slower rate of motion and not by variations in proprioceptive feedback attributed to the loading conditions. In order to test for this possibility one temporal variability measure (Score 1) was re-examined by the method of analysis of covariance which makes it possible to control for movement time. No substantial changes in *F* ratios resulted, and all four of the *F*'s remained significant at the .01 level of confidence. It is concluded, therefore, that the mechanical effect which tends to reduce the average speed of motion is not directly responsible for the decreased variability in timing obtained with the conditions of added damping and mass.

The effect of spring loading observed in the present study lends some support to the hypothesis that the spring should improve spatial accuracy, but the effects for this physical parameter are generally inconclusive since the use of a spring-loaded control and variations in the magnitude of the spring constant did not consistently reduce variability of spatial aspect of movement for either the triangular or the circular movements. Several interpretations of these results for the spring constant are possible. It may be true that spring loading does not lead to significant improvement in position discriminations under any circumstances. Or improvement may be obtained only with extended practice and knowledge of results. Since the present results are confined to a few test trials without knowledge of results and were

preceded by only a few practice trials, they do not provide an answer to the above alternative interpretations. However, a subsequent study (1), which investigated the effects of spring loading upon terminal accuracy of simple rotary arm motions, has demonstrated the effectiveness of this type of loading when practice is extended and knowledge of results are given.

Summary

This study tested the hypotheses that elasticity added to a control should lead to improved spatial accuracy of movements, while damping and mass added to a control should lead to improved temporal accuracy of movement. These hypotheses are derived from the fact that the physical parameters of spring stiffness, damping, and mass characterizing a control determine proportionality relations between applied torque and changes in the position, rate of motion, and acceleration, respectively, of the control. This proportionality provides a basis for learning to associate a given torque with a given characteristic of the movement.

To test these hypotheses *S*s performed simple circular and triangular control motions with a joystick when the control was loaded with various degrees of spring stiffness, or damping, or mass. Three levels of each control parameter were equated originally by means of a psychophysical procedure so that movements opposed by each of the three physical factors were judged by *S*s to involve equal "effort."

In each control loading condition the movements were first practiced for three trials with the help of a visual guide and paced by means of a metronome. The visual and auditory guides were then removed, and *S*s were instructed to reproduce the motions as accurately as possible. Photographic records of all motions were measured for accuracy of temporal and spatial reproductions. It was found that an increase of viscous damping or of inertia of the control resulted in greater uniformity of speed within individual motions, and also in greater uniformity of speed in successive reproductions of the same motion. In the case of the triangular motions increased mass and increased damping led to greater uniformity of peak velocity on each side of the triangle on successive trials. Increased damping and mass also led to slightly lower average velocities of motion, but an analysis of covariance on Score 1 indicated that the reduction in temporal variability was still significant when the effect of movement time was controlled. These results support the hypotheses regarding the utilization of proprioceptive information.

The effect of spring loading upon spatial accuracy of the reproduced motion was in general not significant. It is suggested that extended practice with knowledge of results are needed for effective utilization of cues provided by spring loading. The beneficial effects of mass and viscosity might, of course, also increase with extended training.

References

1. Bahrick, H. P., Bennett, W. F., & Fitts, P. M. Accuracy of positioning responses as a function of spring loading in a control. *J. exp. Psychol.*, 1955, **49**, 437–444.
2. Derwort, A. Ueber die Formen unserer Bewegungen gegen verschiedenartige Widerstaende und ihre Bedeutung fuer die Wahrnehmung von Kraeften. *Z. f. Sinnesphys.*, 1943, **70**, 135–184.
3. Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Ch. 35.
4. Gibbs, C. B. The continuous regulation of skilled responses by kinesthetic feedback. *Brit. J. Psychol.* (Gen. Sec.), 1954, **45**, 25–39.
5. Helson, H. Design of equipment and optimum human operation. *Amer. J. Psychol.*, 1949, **62**, 473–497.
6. Howland, D., & Noble, M. E. The effect of physical constants of a control on tracking performance. *J. exp. Psychol.*, 1953, **46**, 353–360.
7. Jenkins, W. O. A psychophysical investigation of ability to reproduce pressure. In P. M. Fitts (Ed.), *Psychological research on equipment design*. Washington, D. C.: U. S. Govt. Print. Off., 1947.
8. Klemmer, E. T. The discrimination and control of force and velocity in elbow flexion. Unpublished doctor's dissertation, Columbia University, 1952.
9. Weiss, B. The role of proprioceptive feedback in positioning responses. *J. exp. Psychol.*, 1954, **47**, 215–224.

(Received for early publication January 20, 1955)

C36